Synthetic IV Estimation in Panels[†]

Ahmet Gulek Jaume Vives-i-Bastida

Job Market Paper

[Most recent version]

Abstract

We propose a Synthetic Instrumental Variables (SIV) estimator for panel data that combines the strengths of instrumental variables and synthetic controls to address unmeasured confounding. We derive conditions under which SIV is consistent and asymptotically normal, even when the standard IV estimator is not. Motivated by the finite sample properties of our estimator, we introduce an ensemble estimator that simultaneously addresses multiple sources of bias and provide a permutation-based inference procedure. We demonstrate the effectiveness of our methods through a calibrated simulation exercise, two shift-share empirical applications, and an application in digital economics that includes both observational data and data from a randomized control trial. In our primary empirical application, we examine the impact of the Syrian refugee crisis on Turkish labor markets. Here, the SIV estimator reveals significant effects that the standard IV does not capture. Similarly, in our digital economics application, the SIV estimator successfully recovers the experimental estimates, whereas the standard IV does not.

[†]MIT Economics. Contact: vives@mit.edu. Jaume Vives-i-Bastida is especially indebted to Alberto Abadie, Anna Mikusheva and Tobias Salz for their guidance and support. We are also thankful to Daron Acemoglu, Isaiah Andrews, Josh Angrist, Dmitry Arkhangelsky, David Autor, Marinho Bertanha, Clément de Chaisemartin, Jeroen Dalderop, Avi Feller, Amy Finkelstein, Bruno Ferman, Sergi Jiménez-Martín, Guido Imbens, Sylvia Klosin, Max Kasy, Evan Munro, Whitney Newey, Benjamin Pugsley, David Ritzwoller, Ashesh Rambachan, Jann Spiess, Eric Tchetgen Tchetgen, Davide Viviano, Yinchu Zhu and participants in the International Association for Applied Econometrics Annual Conference, the North American and European meetings of the Econometric Society, the BSE Economics Forum in Micro-econometrics, the American Causal Inference Conference, and the Notre Dame Econometrics Junior Conference, among other conferences and seminars, for valuable comments and suggestions. We are grateful to the La Caixa Foundation Graduate Fellowship, the Meta Research PhD Fellowship, the Jerry A. Hausman Fellowship Fund, and the UniCredit Foundation for funding. A previous version of the paper was circulated under the title "Synthetic Instruments in DiD Designs with Unmeasured Confounding."

1. Introduction

In this paper, we propose a Synthetic Instrumental Variables (SIV) estimator that combines instrumental variables and synthetic controls to account for bias due to unobserved confounding. We are interested in panel data settings in which an intervention affects a set of units over time, but we are worried about endogeneity concerns such as the intervention affecting units selectively or differential trends amongst units that received different doses of the treatment. In this context, researchers may turn to differences-in-differences (DiD) designs (Card and Krueger, 2000) or synthetic control (Abadie and Gardeazabal, 2003) designs (SC) in which control units are used to evaluate the counterfactual in absence of the intervention. While these approaches may address part of the endogeneity problems, often valid control units may not exist, as all units may be treated, or control units and treated units may not follow similar paths, violating the parallel trends assumption. Faced with this challenge, we may consider an instrumental variable (IV) approach in combination with the DiD design (for example using a shift-share instrument, e.g. Jaeger et al. (2018a)). In practice, however, the endogeneity concerns may persist as the instrument may be correlated with unobserved confounders in the outcome of interest. The SIV estimator provides a solution to this problem.

To understand the relevance of our setting, consider the use of shift-share instrumental variables (SSIV) to identify causal effects of a treatment or policy by comparing groups or regions more and less exposed to the treatment. Influential examples include studies of the effects of immigration (Card, 2001) and trade (Autor et al., 2013) on labor markets. In this paper, our main empirical application concerns the study of the effect of immigration on Turkish labor markets using the Syrian civil war as an exogenous shock. An intuitive empirical strategy to address the endogeneity of immigrants' location choice, is to use a shift-share design where the distance to the border is the "share" and the aggregate inflow of immigrants is the "shift." Identification in this context relies on regions close to and away from the border to follow parallel trends absent migration flows. The problem with such a design is that regions close and away from the border may be on different economic trajectories before the Syrian civil war starts. These differential trends may bias our estimates of the effect of the refugees on local labor outcomes. Our proposed method, the SIV, creates a synthetic control unit for each region in the pre-intervention period and then debiases the outcomes of interest to account for the differential trends and correct the bias in the two-stage least square estimator (TSLS).

We motivate our method theoretically by deriving consistency and asymptotic normality results in triangular panel designs with unmeasured confounding. We assume that the unobserved error term has two components: an idiosyncratic component that is orthogonal to the instrument and an unobserved heterogeneity component that follows a factor structure. If we could control for the unobserved factor structure the TSLS would be consistent, but we cannot do so directly. Our solution, the SIV, proposes synthetic controls as a way to proxy for the unobserved confounding through interpolation. Under signal-to-noise restrictions and weak primitive assumptions we show that the synthetic IV is consistent and asymptotically normal when the number of units and time periods is large. Through finite sample bounds we highlight that the proposed estimator might be especially sensitive to the noise level and the weakness of the instrument. To guard against small sample biases of the estimator, we propose empirical checks researchers might want to implement in practice, as well as a "doubly robust" ensemble estimator that combines the synthetic IV with a projected synthetic IV that partials out the noise. We also provide an alternative permutation based inference procedure that is exactly valid in small samples.

We show the applicability of our method in a calibrated simulation exercise, by studying the Syrian refugee crisis example, by re-visiting the effect of Chinese imports on US manufacturing employment (Autor et al. (2013)) and by studying the effect of producer rankings on sales in digital platforms. The simulation study shows that the synthetic IV and ensemble estimators outperform the TSLS (with two-way fixed effects) in a variety of settings. Furthermore, the SIV exhibits close to zero bias in settings with moderate and small levels of noise and unmeasured confounding, and the ensemble estimator is shown to be robust in settings with higher noise levels. In a study of the coverage of the synthetic IV estimator confidence intervals we find that it is good in cases in which the estimator exhibits small bias. Following the theoretical properties and the observed behavior under simulations we recommend that researchers implement four checks (in the spirit of the best practices detailed in Abadie and Vives-i-Bastida (2022)) when using the estimator: (1) ensuring that the instrument is not weak after the debiasing, (2) making sure that the estimator achieves good fit in the pre-treatment period, (3) implementing a back test to ensure the good fit is not due to over-fitting to the idiosyncratic noise and (4) ensuring the synthetic controls weights are dense, with no one unit receiving all the weight.

In our study of the effects of Syrian migrants on Turkish labor markets we find that regions close to and away from the border follow different trajectories in the pre-period, potentially biasing the estimates from a shift-share design. However, the SIV corrects for this problem and the debiased estimates do not exhibit pre-trends. Moreover, using SIV leads to different conclusions, relative to the standard shift-share IV, about the effect of immigration in the Turkish context. While the shift-share IV estimator cannot reject that there is no effect of immigration on natives' salaried employment, the synthetic IV estimator finds a statistically significant negative effect. For example, using SSIV we find that a 1 percentage point (pp) increase in refugee/native ratio is associated with a 0.01 pp *increase* in native salaried employment for low-skilled men, whereas using SIV we find that it causes a, statistically significant, 0.16 pp *decrease*. This implies that for every 100 immigrants that arrived to Turkey, 16 low-skilled natives lost salaried jobs. These economically and statistically significant differences between the SSIV and SIV estimates highlight the role of unobserved confounders in the long-standing debate about the labor market effects of immigrants (Borjas, 2017; Peri and Yasenov, 2019).

In our re-analysis of the effect of Chinese imports on US manufacturing, we follow the identification strategy of Autor et al. (2013). We compare regions that were more exposed to Chinese imports based on their pre-existing industrial composition with regions that were less exposed. We first show that the "shares" are correlated with regional growth rates in the pre-period. Regions that were more exposed to the China shock starting from 1990 grew less in 1970s and 1980s. In fact, the difference in growth rates in the 1970s and 1980s is almost identical to the difference in growth rates after the China shock in 1990s and 2000s. The SIV estimator corrects for this pre-trend, and finds smaller effects in the 1990s than the original SSIV, but similar effects in the 2000s. This evidence implies that regional trends between 1970–1990 account for about half of the effect the standard SSIV captures in 1990s. This is intuitive as the economic trajectories between 1970–1990 are more likely to continue in 1990s than 2000s. Our findings contribute to the growing literature estimating the China shock effect under different modelling assumptions (Goldsmith-Pinkham et al., 2020).

Finally, we apply the SIV to study an important question in digital economics: how much do rankings affect producer outcomes? In the context of a large food delivery platform we use preferential contracts given to producers that mechanically increase their ranks in the consumer search wall as an IV for their rank. In this context, we have at our disposal an A/B test in which rank was randomized which allows us to benchmark our observational estimates on the effect rank. As expected, we find that the standard IV (with two-way fixed effects) exhibits positive omitted variable bias (relative to the A/B test) as producers that receive the preferential contracts are in an upwards trend relative to others. The synthetic IV estimates however do not exhibit such bias and recover the A/B test estimates. This examples corroborates the usefulness of our proposed estimator in dealing with unmeasured confounding in IV-DiD settings and provides a new strategy to measure the causal effect of rank in digital platforms using instrumental variables. Given the challenges associated with IVs in this context (as discussed by Rutz et al. (2012)) we see this as a contribution to the literature in digital economics.

This paper contributes to several strands of the literature. First, it complements the growing body of work on addressing unobserved confounding and 'pre-trends' in panel data settings by providing a new method for the IV DiD case. Research in this area is built upon synthetic control based methods (Abadie et al., 2010, 2015; Ben-Michael et al., 2021; Imbens and Viviano, 2023), more general weighting methods such as the synthetic differences in differences (Arkhangelsky et al., 2021a), as well as balancing methods (Hainmueller, 2012), matrix completion methods (Agarwal et al., 2021; Athey et al., 2021) and factor model methods (Anatolyev and Mikusheva, 2022; Bai, 2009). Similarly, our paper complements related work on addressing and evaluating pre-trends in event-study designs, including Freyaldenhoven et al. (2019), Borusyak et al. (2023), Roth (2022) and Ham and Miratrix (2022) among others. A more closely related paper is Arkhangelsky and Korovkin (2023) which provide a novel weighting algorithm to address unobserved confounding in settings in which the exogenous variation comes from aggregate time series shocks. The authors propose a robust estimator that corrects the TSLS bias when the instrument has a product form and there are unobserved aggregate shocks that may affect different units differently. We see our method as complementary to Arkhangelsky and Korovkin (2023), and note that we consider a different setting in which the instrument need not have a product structure and the exogenous variation may come from the time or unit components.

Second, this paper is related to a growing literature studying and relaxing the identification assumptions embedded in shift-share designs. Goldsmith-Pinkham et al. (2020) show that the identification assumptions in SSIV designs are often based on the exogeneity of shares. Borusyak et al. (2022) relax this assumption and provide a framework in which identification can also come from the exogeneity of shifts, allowing shares to be endogenous. Adao et al. (2019) highlight an inference problem that arises from cross-regional correlation in the regression residuals due to similarity of sectoral shares in the US. In the immigration context, Jaeger et al. (2018b) show that past-settlement instruments in practice conflate both short-term and long-term adjustments to immigration shocks, which invalidates the exogeneity assumption. Our method provides an additional tool applied researchers can rely on to address unobserved confounders in the SSIV designs.

Lastly, our main empirical example is related to a large literature studying the effects of immigration using refugee shocks (Card, 1990; Hunt, 1992; Friedberg, 2001; Angrist and Kugler, 2003; Lebow, 2022). More specifically, our focus on the effects of Syrian refugees on Turkish natives and the presence of unobserved confounders in Turkey follows Gulek (2023). Whereas he focuses on the effects on the formal and informal labor markets, we focus on the overall impact on salaried employment and consider heterogeneity across men and women.

The paper proceeds as follows. Section 2 describes the setting and an empirical example. Section 3 presents the synthetic IV estimator and two additional estimators. Section 4 discusses the theoretical results. Section 5 regards extensions of the estimator and inference procedures. Section 6 concerns the simulation study, and section 7 details our three empirical applications.

2. General setting and empirical motivation

We are interested in a panel data setting in which some units of interest are exposed to a (potentially continuous) treatment and there are endogeneity concerns. The researcher may be worried about using a differences-in-differences design as the parallel trends assumption might not hold, but has access to an instrument that *partially* addresses the endogeneity concerns. More precisely, we consider J units indexed by $i = 1, \ldots, J$ that are observed for T periods of time with outcomes of interest Y_{it} and potential outcomes denoted by $Y_{it}(R_{it})$ for a random variable $R_{it} \in \mathbb{R}$. Throughout the paper we assume that the potential outcomes are generated as described by the following assumption.

Assumption 1 (Design). *Outcomes follow*

$$Y_{it}(R_{it}) = \theta R_{it} + U_{it} + \epsilon_{it}$$

where $Y_{it}, R_{it},$ and Z_{it} are observed and $U_{it} = \mu_i' F_t$, for a $k \times 1$ vector of factor loadings μ_i and a vector of common factors F_t, and ϵ_{it} are unobserved. The treatment R_{it} follows

$$R_{it} = \gamma Z_{it} + A_{it} + \eta_{it},$$

where Z_{it} is an instrument satisfying that $Z_{it} = 0$ for $t \leq T_0$, A_{it} is an unobserved heterogeneity term and η_{it} is an idiosyncratic shock potentially correlated with the error term ϵ_{it}.

The main feature of the panel triangular design we consider in Assumption 1 is that the instrument Z_{it} is not active for $t \leq T_0$ and the unobserved components are additively separable into unobserved heterogeneity components (U_{it} and A_{it}) and idiosyncratic error components (ϵ_{it} and η_{it}). The design reflects settings in which an observed intervention starts at T_0 and is used as an instrument or to construct an instrument Z_{it}. For example, shift-share designs satisfy this setting as $Z_{it} = Z_i' H_t$ for shares Z_i and shifts H_t with $H_t = 0$ for $t \leq T_0$. This will be the case for our application to the China shock study. Panel instrumental variable designs in which the instrument becomes active after T_0 (i.e. $Z_{it} = 0$ for $t \leq T_0$) also satisfy this setting, and this will be the case of our digital economics application. A special instance of our design is $R_{it} = 0$ for $t \leq T_0$ (i.e. $A_{it} = \eta_{it} = 0$ for $t \leq T_0$), which is satisfied in our main empirical application to the Syrian refugee crisis, as well as for common frameworks considered in the literature of IV-DID.

The parameter of interest is the expected marginal effect of the treatment R_{it} on the outcome Y_{it},

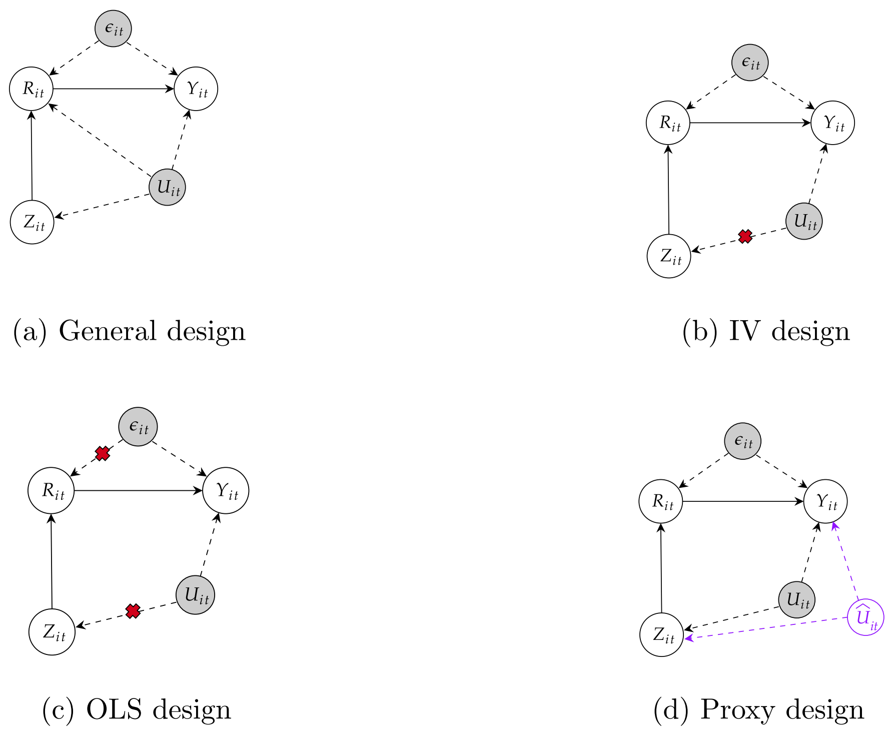
$$\theta = \mathbb{E}\left[\frac{\partial Y_{it}(R_{it})}{\partial R_{it}}\right] = \frac{\partial Y_{it}(R_{it})}{\partial R_{it}}.$$

To understand the potential problems arising in the estimation of θ; it is useful to consider the possible identifying assumptions researchers may posit. To this end Figure 1 describes different assumptions on our design encoded in directed graphs. In this paper, we relax the standard IV independence assumption by allowing the instrument to be correlated with the unobserved term U_{it} as described in panel (a) of Figure 1 and Assumption 2.

Assumption 2 (Partial instrument exogeneity)**.** *The following independence conditions hold* $\epsilon_{it}, \eta_{it}, A_{it} \perp Z_{it}$.

Tu put in context Assumption 2 we consider alternative assumptions researchers may posit. Researchers may consider the independence assumption $R_{it} \perp \epsilon_{it}, U_{it}$ (panel (c) in Figure 1) which in our design is satisfied when $A_{it}, Z_{it}, \eta_{it} \perp \epsilon_{it}, U_{it}$ pairwise. This assumption is not implied by Assumption 2, but if it holds, then the OLS estimator of θ is unbiased and the researcher could recover θ by regressing Y on R. In many empirical settings, however, R_{it} is likely correlated with the unobserved components. For example, in immigration settings refugees might take into account local labor market conditions and trends when choosing where to re-locate or alternatively may relocate based on geographical distance. Researchers may therefore rely on instrument Z_{it} to address this concern. Common instruments in the immigration literature to address location choice endogeneity include past-settlement indicators or travel distance (Card, 2001; Angrist and Kugler, 2003).

Figure 1: Triangular designs



(a) General design

(b) IV design

(c) OLS design

(d) Proxy design

Notes: Directed acyclic graphs representing the independence assumptions implicit for different designs. Variables shaded grey are unobserved.

A valid instrument requires that (1) the only channel affecting the outcome Y is through the treatment R (the exclusion restriction) and (2) that the instrument is as good as randomly assigned. In our design, this would require that $\epsilon_{it}, \eta_{it}, U_{it}, A_{it} \perp Z_{it}$, which is also not implied by Assumption 2, as reflected in panel (b) in Figure 1. We are interested in cases in which the instrument Z_{it} is not perfectly valid due to failing to be as good as randomly assigned. This may be the case in some relevant empirical settings. For instance, in immigration examples regions that received immigrants in the past or were closer to the immigrants' origin may follow different trends than other regions. Given that models of migrant location decisions (Llull (2017), Bartel (1989)) often involve agents taking expectations over both settling costs and the economic returns of settling in a particular region; it is likely that a single instrument Z_{it} may not simultaneously address both sets of endogenous variables. However, we may believe that the instrument Z_{it} is valid for one set of variables (settling costs for example) and that the design would be valid if we could control for economic trends.

In other words, the researcher may believe there exists an omitted variable that is correlated with the instrument and the outcome of interest. This motivates us to relax the instrument independence conditions as posited in Assumption 2 to distinguish between the unobserved component ϵ_{it} unrelated to the instrument from the unobserved heterogeneity component U_{it}.

Assumptions 1 and 2 imply that the instrument Z_{it} correctly addresses the endogeneity problem due to the unobserved component ϵ_{it}, but not the omitted variable bias due to U_{it}. If we could observe U_{it}, we would control for it and the IV design would be valid. However, U_{it} is unmeasured and, therefore, the statistical problem we consider is that of finding a valid proxy control $\hat{U}_{it}$ for U_{it} (as depicted in panel (d) of Figure 1). Different approaches have been considered in the literature to tackle this problem. If additional variables are available, strategies have been proposed to combine the observed variables and the instrument to proxy for U_{it} directly (see Miao et al. (2018) and Deaner (2021)). If there exists a donor pool of units never exposed to the treatment R, researchers may opt for a different design and use synthetic controls to partial out U_{it} (Cengiz and Tekgüç, 2022). However, in many empirical settings additional variables or additional control units are not available, and in these cases researchers often rely on parametric assumptions to control for U_{it}. Common examples used in the literature include two-way fixed effects, $U_{it} = \alpha_i + \delta_t$, linear time trends (Wolfers, 2006), $U_{it} = \alpha_i \times t$, or grouped region time fixed effects (Stephens Jr and Yang, 2014), $U_{it} = \alpha_g \delta_t$ for region groups g. The choice of parametric form is often driven by domain knowledge, or, when possible, by showing that the instrument Z_{it} is not correlated with the outcome in the pre-treatment period (see Danieli et al. (2024) for a review of IV falsification tests).

In this paper, we propose a strategy to directly use the pre-treatment period ($t \leq T_0$) to flexibly control for U_{it} without estimating a specific functional form. Our proposed method, the *synthetic* IV, uses the pre-treatment period to estimate synthetic control weights that interpolate across units to partial out U_{it} in the post-treatment period ($t > T_0$). The idea of using the pre-treatment period as a way to address an omitted variable bias in panel IV estimates is not completely novel, and this is often done in IV-DID settings by instrumenting the difference of outcomes before and after treatment (sometimes called the differences in Wald estimator). However, the validity of this approach still relies on strong parametric assumptions on U_{it}, for example Danieli et al. (2024) show for that for a binary setting under the parallel trends assumption (which implicitly imposes that $U_{it} = \alpha_i + \delta_t$), the differences in Wald estimator is an unbiased estimator for θ. The method we propose works for continuous treatments and instruments, and is valid under more general functional forms; linear factor models $U_{it} = \mu_i' F_t$. Given that many economic trend variables can be described through an interactive factor structure we believe that our method can be quite useful to researchers to flexibly control for unmeasured confounding and prevent the specification search over different fixed effect combinations. Furthermore, linear factor model assumptions are common in the literature for dealing with unmeasured confounding in panel settings. A related paper that also relies on a factor model structure is Arkhangelsky and Korovkin (2023) in which an aggregation scheme based on Arkhangelsky et al. (2021b) is used to control for unmeasured aggregate confounders between a time series instrument Z_t and a panel outcome Y_{it}. We see our paper as complementing the novel work of Arkhangelsky and Korovkin (2023) for cases in which we have a panel instrument Z_{it} and our identifying assumptions may come from unit level variation in the instrument or from time series variation, which include the commonly used shift-share designs.

Example 1 (the Syrian refugee crisis) To further motivate why the setting described under Assumptions 1-2 and in Figure 1 is relevant to applied work, consider our main empirical example: the effect of the Syrian refugee crisis on Turkey's local labor markets. The Syrian civil war started in March 2011 and by 2017, 6 million Syrians had sought shelter outside of Syria with 3.5 million locating in Turkey.[1] Figure 2 panel (a) shows the growth in the number of Syrian refugees in Turkey over time and panel (b) shows the geographic dispersion of the refugees. Given the structure of the Syrian refugee shock a natural approach to estimating the impact of refugees on local labor outcomes is that of a shift-share instrumental variable design that exploits the exogenous time shock of the civil war and the differential impact across units.

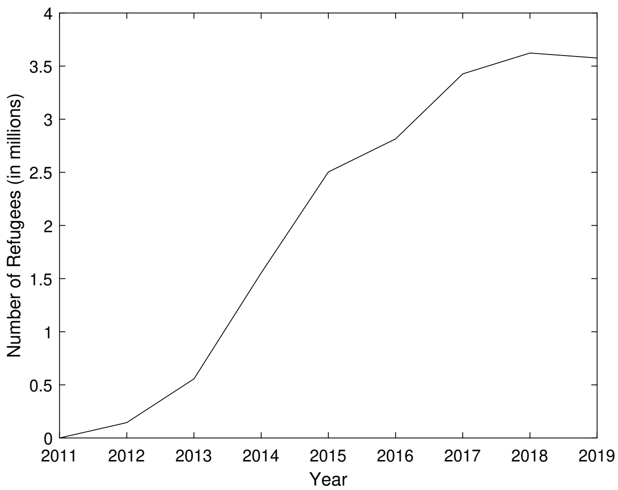
To relate the Syrian example to our setting let R_{it} denote the refugee/native ratio at province-year level and consider a travel distance shift-share instrument, as is common in the mass-immigration literature (Angrist and Kugler, 2003; Aksu et al., 2022).

$$Z_{it} = \underbrace{\bar{H}_t}_{\text{shift}} \times \underbrace{Z_i}_{\text{share}},$$
$$Z_i = \sum_{s=1}^{13} \lambda_s \frac{1}{d_{i,s}}$$

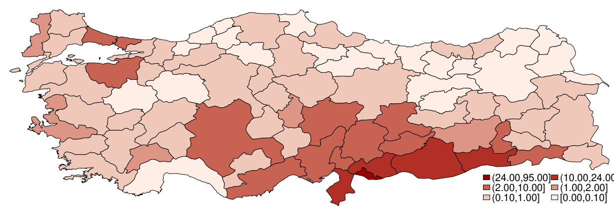
where $\bar{H}_t$ is the number of refugees in Turkey in year t, $d_{i,s}$ is the travel distance between

[1]Turkey hosts the largest number of refugees in the world (UNHCR, 2021).

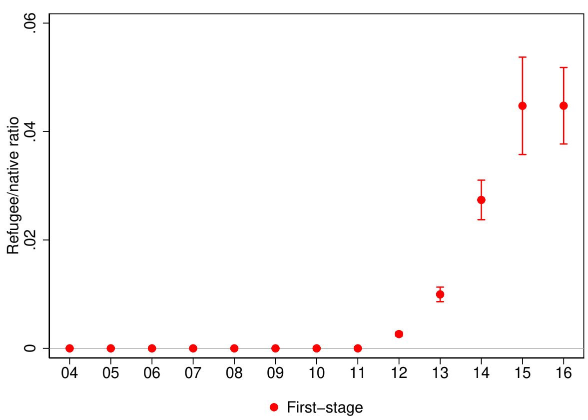
Figure 2: The Syrian refugee shock.



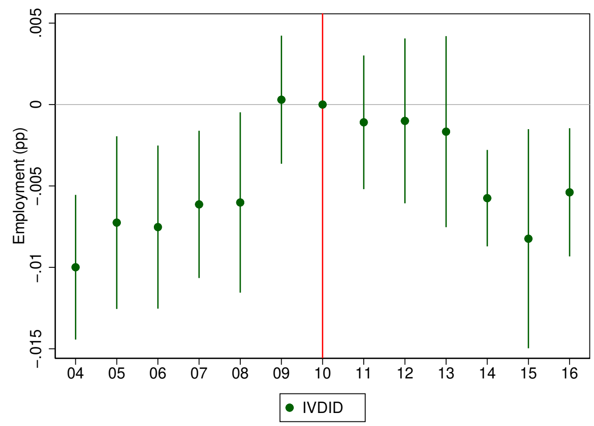
(a) Number of refugees



(b) Refugee shares



(c) First stage



(d) Reduced form

Notes: In event-study designs the 95% confidence intervals are plotted. The F-stat of the main first-stage regression is 154. In Panels (c) and (d) the x axis shows the years 2004-2016 in 2 digit notation.

Turkish region i and Syrian governorate s, λ_s is the weight given to Syrian governorate s which we set it be proportional to the population share of s.[2] In panel (c) of Figure 2 we plot the first stage coefficients interacted with time dummies from the TWFE specification that is commonly estimated in the literature

$$R_{it} = \sum_{k \neq 2010} \gamma_k (\mathbb{1}\{t = k\} \times Z_i) + \alpha_i + \delta_t + \eta_{it}.$$

The first stage regression tests whether the instrument predicts refugees' location choice every year. As expected, distance is a strong predictor of the refugee treatment intensity. The F-stat of the shift-share first-stage (where we regress R_{it} on Z_{it} while controlling for region and time f.e.) is 154. The problem arises when one considers the reduced form of local wage-employment (salaried employment) of the natives that did not finish high-school (low-skill)[3] on the instrument

$$Y_{it} = \sum_{k \neq 2010} \theta_k (\mathbb{1}\{t = k\} \times Z_i) + \alpha_i + \delta_t + \epsilon_{it}, \tag{1}$$

which is displayed in panel (d) of Figure 2. Between 2004–2010 (before the refugee crisis began), the provinces closer to the border observed employment gains compared to other regions. Being one standard deviation closer to the border predicts a wage-employment growth of 1 pp between 2004 and 2009. Given that the regions that are predicted by the instrument to receive immigrants were following different trends *before* the shock, it is likely that the IV-DID design does not satisfy the parallel trends assumption implicit in the TWFE specification. This suggests that there exists an unmeasured confounder U_{it} in the Syrian crisis empirical setting. The appearance of pre-trends in similar designs is a common problem in practice (Wolfers, 2006; Stephens Jr and Yang, 2014; Gulek, 2023) and has been discussed extensively in the literature (Roth, 2022; Freyaldenhoven et al., 2019). While our main design described by Assumption 1 considers a common, time-invariant, parameter θ, in the Appendix (section 1.7) we extend it to a time-varying event study design like the one considered in this example.

[2] The idea is that all else equal, more Syrians would be expected to come from the more populous regions.

[3] This is the key outcome of interest because Syrian refugees were substantially less educated compared to the Turkish population, and hence constitute largely a low-skill immigration shock. We provide more details about the setting in the Appendix.

Example 2 (the China shock) Another shift-share research design that highlights the problem of interest is used by Autor et al. (2013) to evaluate the effect of Chinese exports on manufacturing employment in the US. The authors are interested in the effect of US import's exposure to China on the growth in percent manufacturing employment (Y_{it}) for a US commuting zone i during a decade t. Given that import exposure is endogenous, Autor et al. (2013) instrument it by the increase in Chinese imports by high-income countries (Z_{it}). It is important to note that the instrument has a shift-share structure. Therefore, we can interact the exposure components in the instrument with time indicators to see whether they predict changes in the outcome *before* the shock occurs as we did for the Syrian example. In particular, we estimate the following reduced form regression

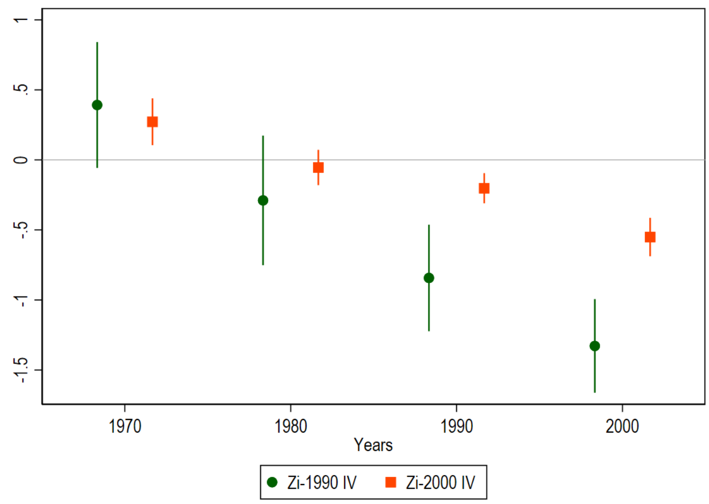
$$Y_{it} = \sum_k (\bar{Z}_{i,h} \times \mathbb{1}\{t = k\})\beta_{k,h} + \delta_t + \epsilon_{it} \tag{2}$$

where δ_t is a time fixed effect, ϵ_{it} is an error term, and $\beta_{k,h}$ with $h = 1990, 2000$ are the event-study estimates of interest for the two exposure measures $\bar{Z}_{i,h}$ considered by the authors in constructing the instrument Z_{it}. In section 7.2 we explain in detail how the shift-share instrument is constructed and replicated the main tables of Autor et al. (2013). Note, that we do not include region fixed effects following the original paper, but in principle we could at the cost of having to normalize one of the pre-period estimates to zero.

Figure 3 shows the event studies estimates for the reduced form regression (2). Each data point represents a coefficient estimate for a given decade, with the intervention (the China shock) occurring in the decade of 1990-2000. As it can be seen, for both exposure measures it appears that before the intervention regions more affected by Chinese imports where potentially on a downward trend in terms of manufacturing employment, raising concerns over the exogeneity of shares assumption implicit in the shift-share design. This example highlights that unmeasured confounding may be a common problem in shift-share designs. Autor et al. (2013) partially address these concerns by using additional covariates to control for the differential trends. In section 7.2 we revisit this example and show how our proposed method can correct for the unmeasured confounding without relying on additional data.

In the following section we describe our proposed solution, the synthetic IV estimator, that flexibly controls for the unmeasured confounder U_{it}.

Figure 3: Reduced-form estimates for the China shock



Notes: Event-study estimates for (2) using the exposure components $\bar{Z}_{i,1990}$ and $\bar{Z}_{i,2000}$. The time periods consist of four decades 1970-1980, 1980-1990, 1990-2000 and 2000-2007, with the intervention (the China shock) occurring in the decade of 1990-2000.

3. The synthetic estimator

The synthetic estimator consists of two steps. In the first step we find synthetic controls for each unit in a pre-period ($t < T_0$) and generate counterfactual estimates for Y_{it}, R_{it} and Z_{it} for a post period. In the second step, as in the standard IV estimator, we use these counterfactual estimates to compute the first stage and reduced form estimates. To describe the procedure, consider J units indexed by $j = 1, \ldots, J$ observed for T periods of time. We are interested in an outcome of interest Y_{it} with potential outcomes $Y_{it}(R_{it})$ indexed by random variable R_{it}.

Step 1: for each $j \in \{1, \ldots, J\}$ we find the synthetic control weights $\hat{w}_j^{SC}$ by solving the following program for the pre-treatment period $t \in \{1, \ldots, T_0\}$

$$\hat{w}_j^{SC} \in \text{argmin}_{w \in \mathcal{W}} \|D_j^{T_0} - D_{-j}^{T_0}{}'w\|^2, \tag{3}$$

for

$$\mathcal{W} = \{w \in \mathbb{R}^J \mid \|w\|_1 \leq C\}, \tag{4}$$

where $C \in (0, \infty)$ is a regularization hyper-parameter, D^{T_0} is the $J \times p$ design matrix that includes pre-treatment outcomes Y_{jt} and treatments R_{it} for $t < T_0$ where $p = 2T_0$, with

$D_j^{T_0}$ denoting the predictors for unit j and $D_{-j}^{T_0}$ the $(J-1) \times p$ matrix of predictors for the other units.[4] The intuition for matching outcomes and treatments is that the finite sample behavior of the estimator will depend on the pre-treatment fit of R_{it} and Y_{it}. In the case in which R_{it} is not present in the pre-period, as is the case of the Syrian refugee example (2), the design matrix includes only the pre-treatment outcomes, such that $D^{T_0} = Y^{T_0}$. The l_1 norm constraint on the weights ensures that there is some amount of regularization. This program is a relaxation of the standard synthetic control objective, sometimes called the constrained lasso (Doudchenko and Imbens, 2016). In our empirical applications, we compute the weights using the standard synthetic control restriction that the weights are in the simplex (i.e. $\mathcal{W} = \{w \mid w_j \geq 0, \quad \sum_j w_j = 1\}$). Our theoretical results will be valid for this case when $C = 1$ is chosen in solving program (3). We find in the simulations, and empirical exercises, that the additional regularization provided by the simplex restrictions offers good finite sample performance.

Once the synthetic control weights are computed, we define the following quantities for all t in $1, \ldots, T$

$$\hat{Y}_{it}^{SC} = \sum_{j \neq i} \hat{w}_{ij}^{SC} Y_{jt},$$
$$\hat{R}_{it}^{SC} = \sum_{j \neq i} \hat{w}_{ij}^{SC} R_{jt},$$
$$\hat{Z}_{it}^{SC} = \sum_{j \neq i} \hat{w}_{ij}^{SC} Z_{jt},$$

which we label the synthetic outcome, treatment level and instrument respectively. Then, we define the *debiased* values for $t > T_0$ as the difference between the observed values and the synthetic values

$$\tilde{Y}_{it} = Y_{it} - \hat{Y}_{it}^{SC},$$
$$\tilde{R}_{it} = R_{it} - \hat{R}_{it}^{SC},$$
$$\tilde{Z}_{it} = Z_{it} - \hat{Z}_{it}^{SC}.$$

Step 2: Given $\{\tilde{Y}_{it}, \tilde{R}_{it}, \tilde{Z}_{it}\}_{t=T_0+1}^{T}$, we estimate the first stage and reduced form by pooled

[4] Researchers may choose to weight the columns of the design matrix D^{T_0} according to different weights, as proposed by Abadie et al. (2010).

OLS regression

$$
\begin{aligned}
\tilde{\pi} &\in \arg\min_{\pi} (\tilde{Y} - \tilde{Z}\pi)'(\tilde{Y} - \tilde{Z}\pi), \\
\tilde{\beta}_1 &\in \arg\min_{\beta} (\tilde{R} - \tilde{Z}\beta)'(\tilde{R} - \tilde{Z}\beta).
\end{aligned}
$$

where $\tilde{Y} = vec(\tilde{Y}^T)$, $\tilde{R} = vec(\tilde{R}^T)$ and $\tilde{Z} = vec(\tilde{Z}^T)$ are the $(J(T - T_0) \times 1$ vectors of the debiased values. Then, the estimated average marginal effect is given by the standard IV estimate

$$
\tilde{\theta}^{SIV} = \frac{\tilde{\pi}}{\tilde{\beta}_1},
$$

which we denote the *synthetic* IV estimator. Given that our framework is a just-identified IV design, the IV estimator is equivalent to the two-stage least-squared estimator (TSLS) given by

$$
\tilde{\theta}^{TSLS} = \left(\sum_{it} \tilde{Z}_{it} \tilde{R}_{it} \right)^{-1} \sum_{it} \tilde{Z}_{it} \tilde{Y}_{it}.
$$

Our main asymptotic results are also valid for the estimator that uses the instrument Z_{it} instead of the de-biased instrument $\tilde{Z}_{it}$

$$
\tilde{\theta}^{TSLS}_{YR} = \left(\sum_{it} Z_{it} \tilde{R}_{it} \right)^{-1} \sum_{it} Z_{it} \tilde{Y}_{it},
$$

and the estimator that only debiases the instrument Z_{it}

$$
\tilde{\theta}^{TSLS}_{Z} = \left(\sum_{it} \tilde{Z}_{it} R_{it} \right)^{-1} \sum_{it} \tilde{Z}_{it} Y_{it}.
$$

In the theory and simulation sections we show that while these estimators are similar to the proposed *synthetic* IV estimator (SIV), they may have worse finite sample properties. To understand the differences between the standard IV and the SIV, we expand the *debiased* variable $\tilde{Y}_{it}$ in terms of $\tilde{R}_{it}$

$$\begin{aligned}\tilde{Y}_{it} &= Y_{it} - \hat{Y}_{it}^{SC} \\ &= \theta R_{it} + \mu_i' F_t + \epsilon_{it} - \sum_{j \neq i} \hat{w}_{ij}^{SC} Y_{jt} \\ &= \theta \tilde{R}_{it} + (\mu_i - \sum_{j \neq i} \hat{w}_{ij}^{SC} \mu_j)' F_t + \epsilon_{it} - \sum_{j \neq i} \hat{w}_{ij}^{SC} \epsilon_{jt}.\end{aligned} \tag{5}$$

It follows that the *synthetic* IV estimator for the regression of $\tilde{Y}$ on $\tilde{R}$ instrumented by $\tilde{Z}$ for $t > T_0$ recuperates the true parameter θ up to two potential bias terms.

$$\begin{aligned}\tilde{\theta}^{TSLS} &= \left(\sum_{i,t>T_0} \tilde{Z}_{it} \tilde{R}_{it} \right)^{-1} \sum_{i,t>T_0} \tilde{Z}_{it} \tilde{Y}_{it} \\ &= \theta + \left(\sum_{i,t>T_0} \tilde{Z}_{it} \tilde{R}_{it} \right)^{-1} \sum_{i,t>T_0} \tilde{Z}_{it} \left(\mu_i - \sum_{j \neq i} \hat{w}_{ij}^{SC} \mu_j \right)' F_t \\ &\quad + \left(\sum_{i,t>T_0} \tilde{Z}_{it} \tilde{R}_{it} \right)^{-1} \sum_{i,t>T_0} \tilde{Z}_{it} \left(\epsilon_{it} - \sum_{j \neq i} \hat{w}_{ij}^{SC} \epsilon_{jt} \right).\end{aligned} \tag{6}$$

Similarly, the standard TSLS estimator $\hat{\theta}^{TSLS}$ can also be decomposed

$$\hat{\theta}^{TSLS} = \theta + \left(\sum_{i,t>T_0} Z_{it} R_{it} \right)^{-1} \sum_{i,t>T_0} Z_{it} \mu_i' F_t + \left(\sum_{i,t>T_0} Z_{it} R_{it} \right)^{-1} \sum_{i,t>T_0} Z_{it} \epsilon_{it}. \tag{7}$$

In both cases, the bias in estimating θ will depend on a term involving the unobserved factor structure $\mu_i' F_t$ and a term involving the idiosyncratic error term ϵ_{it}. Under the partial instrument validity Assumption (2) we might expect the term involving $Z_{it}\epsilon_{it}$ to be close to zero in probability under suitable assumptions, but the term involving $\mu_i' F_t$ will cause omitted variable bias in the IV estimates. The intuition for the SIV estimator is that if the partialling out procedure successfully removes the $\mu_i' F_t$ term, thanks to the synthetic controls matching the factor loading μ_i for each unit, then there is no omitted variable bias. In section 4 we give conditions on the model primitives under which this is the case and the SIV estimator is consistent.

Common synthetic control weights The decompositions (5) and (6) clarify why it is necessary that a common set of synthetic control weights $\hat{w}_i^{SC}$ is used to generate the debiased outcomes $\tilde{Y}$ and treatments $\tilde{R}$. Suppose, that instead different weights w_1 and w_2 were used in constructing $\tilde{Y}_{it}$ and $\tilde{R}_{it}$. Let $\tilde{Y}^w_{it}$, $\tilde{R}^w_{it}$ and Z^w_{it} denote the debiased quantities in which weights w were used. In this case, the SIV estimator has the following decomposition

$$\begin{aligned}\tilde{\theta}^{TSLS}_Z &= \left(\sum_{i,t>T_0} \tilde{Z}_{it}\tilde{R}^{w^2}_{it}\right)^{-1} \sum_{it} \tilde{Z}_{it}\tilde{Y}^{w^1}_{it} \\ &= \theta \left(\sum_{i,t>T_0} \tilde{Z}_{it}\tilde{R}^{w^2}_{it}\right)^{-1} \sum_{it} \tilde{Z}_{it}\tilde{R}^{w^1}_{it} \\ &\quad + \left(\sum_{i,t>T_0} \tilde{Z}_{it}\tilde{R}^{w^2}_{it}\right)^{-1} \sum_{i,t>T_0} \tilde{Z}_{it}\left(\mu_i - \sum_{j\neq i} w^1_{ij}\mu_j\right)' F_t \\ &\quad + \left(\sum_{i,t>T_0} \tilde{Z}_{it}\tilde{R}^{w^2}_{it}\right)^{-1} \sum_{i,t>T_0} \tilde{Z}_{it}\left(\epsilon_{it} - \sum_{j\neq i} w^1_{ij}\epsilon_{jt}\right).\end{aligned} \tag{8}$$

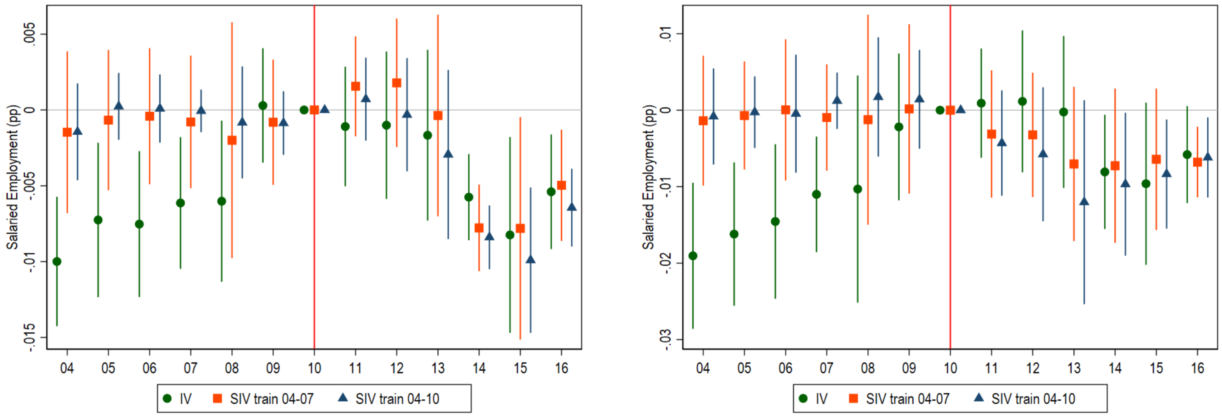
The problem of using different weights is that the SIV estimator will estimate θ up to the term $G = \left(\sum_{i,t>T_0} \tilde{Z}_{it}\tilde{R}^{w^2}_{it}\right)^{-1} \sum_{it} \tilde{Z}_{it}\tilde{R}^{w^1}_{it}$, which unless $w_1 = w_2$ may not be one in finite samples, and, in general, may not converge to one.

The focus of this paper is not to describe what parameters can be identified in this setting in terms of potential outcomes, as we consider the marginal effect defined in section 2. It is possible, however, to derive latent average treatment effect characterization in the case of discrete valued instruments and treatments under a modification of the standard monotonicity assumption. See Mogstad and Torgovitsky (2024) for an in-depth discussion of unobserved heterogeneity in treatment effect in IV models, and for discussion of identification in related IV difference-in-difference settings see Borusyak and Hull (2020). For a discussion of identification of continuous treatment effects in DiD designs we refer readers to de Chaisemartin et al. (2024).

Example 1 (applying the SIV) We show how the synthetic IV estimator works by applying it to the Syrian refugee example. We proceed with the first step by computing the synthetic control weights for each Turkish region by solving problem (3) with $\mathcal{W} = \{w \mid w_j \geq 0, \quad \sum_j w_j = 1\}$.[5] We then compute the debiased variables $\{\tilde{Y}_{it}, \tilde{R}_{it}, \tilde{Z}_{it}\}^T_{t=T_0+1}$ and

[5] As an additional step, we normalize the outcome variable Y before solving for the weights.

Figure 4: Comparing IV and SIV for the Syrian refugee shock.



(a) Salaried employment.

(b) Salaried employment for men.

Notes: panel (a) shows the event-study estimates for regressions 1 and 9 for the shift-share IV and SIV for salaried employment respectively; it replicates panel (d) in Figure 2 for the SIV. Panel (b) shows the same event-study estimates for salaried employment for men. Two SIV estimators are estimated, SIV train 04-10 uses all the pre-treatment periods to compute the synthetic weights in program (3), while SIV train 04-07 only uses the first 4 time periods.

estimate the reduced form regression (1) with the debiased data

$$\tilde{Y}_{it} = \sum_{k \neq 2010} \tilde{\theta}_k (\mathbb{1}\{t = k\} \times \tilde{Z}_i) + \alpha_i + \delta_t + \epsilon_{it}, \tag{9}$$

where $\tilde{Z}_i$ is the debiased instrument share component. In panel(a) of Figure 4 we plot the event study estimates for salaried employment as we did in Figure 2. As we saw before, the IV estimator (green circles) exhibits large pre-trends, but the SIV estimator (blue triangles) does not. To check that indeed the absence of pre-trends for the SIV estimator is not due to over-fitting to the pre-treatment period idiosyncratic noise (ϵ_{it}), we also estimate a backdated SIV in which only a subset of the pre-treatment periods is used in estimating the weights in program (3). The backdated SIV (in orange squares) also shows no pre-trends, providing evidence that the synthetic IV estimator is successfully capturing the unobserved U_{it} term. In the following sections, we highlight the theoretical properties of the SIV estimator and other empirical checks, such backdating, researchers can do to check that robustness of the estimator. In the Appendix (section 1.7), we extend the theoretical results for the SIV estimator to the event study design considered in this example and the empirical applications.

A feature of the synthetic IV estimator is that it accounts for unmeasured confounding that may change over time, and, crucially, differentially before and after the intervention at T_0. This can be seen in Figure 4. In panel (a), while the SIV accounts for the pre-trends before T_0, the estimates after T_0 see changes smaller in magnitude relative to the pre-trend size and non-linear in time, with effects increasing at the start of the post-treatment period and decreasing at the end of the period. If researchers had estimated a linear trend instead ($U_{it} = \alpha_i \times t$) as is common in the literature, the post-treatment estimates would have been shifted downwards significantly. The flexibility of the SIV allows for cases in which the trend in the post-treatment period might be different, for example if we believe Turkish regions close to Syria are in growing path (catching up to richer regions) we might expect the confounding to be smaller in the post-treatment period. Furthermore, the SIV also allows flexibility in functional form across outcomes. In panel (b) of Figure 4 we show the event-study estimates for a different outcome, formal salaried employment for men. While the pre-trends for this outcome are similar than for salaried employment, the SIV post-treatment estimates are shifted downwards uniformly. In section 7.1 we revisit this empirical example and show that SIV can make a real difference compared to the standard IV in estimating the effects of refugees on several local labor outcomes. To preview the results, a researcher using IV would find no effect on natives' or men's salaried employment. Using SIV, however, we find a statistically significant negative effect for both outcomes. We discuss the relevance of these findings in section 7.1.

In sections 4 and 6 we discuss the theoretical properties of the synthetic IV estimator and investigate its finite sample properties through simulations. We highlight that in well behaved settings with low noise but significant correlation between the instrument and the unobserved factor structure the SIV estimator can provide reliable estimates of the true effect. In section 7 we re-visit our three empirical applications using the SIV estimator.

4. Theoretical Results

In this section we provide theoretical guarantees for the SIV estimator. To characterize the behavior of the standard IV and the synthetic IV estimators it is key to understand the behavior of the terms involving the unobserved factor μ_i in decompositions (6) and (7). In order to do so, we impose more structure on the primitives of the design described in assumptions 1 and 2.

Assumption 3 (Model primitives)**.** *Assumptions on the factor structure, the error components and the instruments are as follows.*

- *The common factors are bounded such that for all t, $|F_{lt}| \leq \bar{F}$ for $l = 1, \ldots, k$. Furthermore, the matrix $F_{T_0}F'_{T_0}$ has minimum eigenvalue ξ such that $\xi/T_0 > 0$, where F_{T_0} is the $k \times T_0$ matrix of common factors F_t for $t \leq T_0$. The factor loadings are bounded such that for all i $|\mu_i| \leq c_\mu$.*

- *The unobserved term A_{it} is bounded such that for all i, t $|A_{it}| \leq c_A$. Furthermore, $\frac{1}{JT_1}\sum_{i,t>T_0} A_{it} \xrightarrow{p} 0$ as $JT_1 \to \infty$.*

- *The instrument $Z_{it} \in \mathcal{Z}$ is bounded such that for all i, t $|Z_{it}| \leq c_z$ and $\frac{1}{JT_1}\sum_{i,t>T_0} Z_{it}^2 \xrightarrow{p} Q_Z > 0$.*

- *The instrument Z_{it} and the unobserved factor structure satisfy*

$$\frac{1}{JT_1} Z' M_U Z \xrightarrow{p} Q > 0,$$

  as $JT_1 \to \infty$ for $T_1 = T - T_0$, where $Z = vec(Z_1^T)$ is a $JT_1 \times 1$ vector of instruments and $M_U = I - U_{JT_1}(U'_{JT_1}U_{JT_1})^{-1}U'_{JT_1}$ is the $JT_1 \times JT_1$ residual maker matrix for $U_{JT_1} = vec(U^{T_1})$. Furthermore, the first stage parameter satisfies $\gamma > 0$.

- *ϵ_{it} and η_{it} are i.i.d mean zero subGaussian random variables with variance σ_ϵ^2 and σ_η^2 respectively, finite covariance $\sigma_{\epsilon\eta} = \mathbb{E}[\epsilon_{it}\eta_{it}]$ and bounded fourth moments.*

Assumption 3 has three parts. First, we assume that the model primitives are bounded. This is a common assumption in papers analyzing the behavior of synthetic control estimators and rules out weak factors. Second, we assume that the instrument is strong and not perfectly correlated with the unobserved factor structure. That is, after projecting out the unobserved confounder U_{it} enough variation remains in the instrument. This requirement avoids weak instrument problems and in the simulation discussion we highlight the importance of this assumption for the finite sample performance of the synthetic IV estimator. Finally, we assume that the unobserved error terms η and ϵ are i.i.d, but potentially correlated. This assumption can be weakened to allow for time series correlation, however in our main results the time series dependence is present through the unobserved factor structure $\mu_i' F_t$.

Observe that under our design and Assumption 3 the term in decomposition (7) and (6) depending on the idiosyncratic shocks will converge to zero in probability. In the appendix, we derive a finite sample bound for this term (see Lemma A.1). On the other hand, the unobserved factor term $\sum_{i,t>T_0} Z_{it}\mu_i' F_t$ need not converge to zero in probability at a $1/(JT_1)$ rate. Therefore, in general the TSLS estimator, with or without fixed effects, will be asymptotically biased. The synthetic IV estimator will also be biased in finite samples as $\sum_{i,t>T_0} \tilde{Z}_{it}\tilde{\mu}_i' F_t$ need not be zero, but the bias will depend on how well the synthetic control procedure in step 1 can partial out the factor loadings $\tilde{\mu}_i$.[6] To give conditions for consistency, we condition on the unobserved factor structures and consider ϵ, η and the instrument Z as the source of randomness in our design. The following result gives a bound on the $\sum_{i,t>T_0} \tilde{Z}_{it}\tilde{\mu}_i' F_t$ in terms of model primitives and the mean-absolute deviation of pre-treatment values of Y and R, defined as $\text{MAD}(\tilde{Y}^{T_0}) = \frac{1}{JT_0}\sum_{i,t\leq T_0} |\tilde{Y}_{jt}|$.

Theorem 1 (Factor term consistency)**.** *Under Assumptions 1-3, for $t > T_0$ conditional on the unobserved components $\mu_i' F_t$ and A_{it}, the following bound holds for all J, T_1 and T_0*

$$\frac{1}{JT_1}\mathbb{E}\left[\left|\sum_{i,t>T_0} \tilde{Z}_{it}\tilde{\mu}_i' F_t\right|\right] \leq \left(\frac{\bar{F}^2 k c_z c}{\underline{\xi}}\right)\left(2c\sqrt{\frac{J}{T_0}}\sigma_\epsilon \right.$$
$$\left. + \mathbb{E}\left[\frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{Y}_{jt}|\right] + \theta\mathbb{E}\left[\frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{R}_{jt}|\right]\right)$$

where $c = 1 + C$ and all other terms are defined in Assumptions 1-3. Furthermore, under the same assumptions, as $JT_1 \to \infty$, $\mathbb{E}MAD(\tilde{Y}^{T_0}) \to 0$, $\mathbb{E}MAD(\tilde{R}^{T_0}) \to 0$, and $\sqrt{\frac{J}{T_0}} \to 0$,

$$\frac{1}{JT_1}\sum_{i,t>T_0} \tilde{Z}_{it}\tilde{\mu}_i' F_t \xrightarrow{p} 0.$$

Theorem 1 states that the bias term that depends on the unobserved factor structure can be bounded above by the expected mean absolute deviation of the outcome variable in the pre-treatment period and a term that depends on the likelihood of pre-treatment "over-fitting". This is a standard bound in papers evaluating the properties of synthetic control estimators (see Abadie et al. (2010) for the first example in the literature and Vives-i-Bastida (2022) for a example with covariates). It highlights the dependence of the estimator on good pre-treatment fit (see Ferman and Pinto (2021) for a discussion of synthetic controls with imperfect pre-treatment fit). In particular, the bound depends on the error noise level σ_ϵ and the ratio $\sqrt{J/T_0}$. In settings, in which the we have a small amount of pre-treatment periods, a large number of units, or in which the noise level is high, perfect interpolation of the noise is more likely, biasing the estimator. A discussion in Abadie and Vives-i-Bastida (2022)

[6] Under weak conditions on the time series (β-mixing, or covariance stationarity), it may also be possible to directly correct this bias using a cross-fitting procedure in the spirit of the method proposed in Chernozhukov et al. (2022) for the synthetic control framework.

highlights the importance of pre-treatment fit and over-fitting for performance of synthetic control estimators through a simulation study. Similarly, we evaluate the performance of the synthetic IV estimator in simulations in section 6 and find that the estimator performs well even in settings with moderate $\sigma_\epsilon\sqrt{J/T_0}$.

To provide conditions under which $\mathbb{E}\mathrm{MAD}(\tilde{Y}^{T_0}) \to 0$ and $\mathbb{E}\mathrm{MAD}(\tilde{R}^{T_0}) \to 0$, we consider a relaxation of rank proposed by Rudelson and Vershynin (2007) that allows for *small* perturbations.

Assumption 4 (Numerical rank assumption)**.**

With probability one, for all J and T_0, the pre-treatment matrices have bounded numerical rank, $\frac{\|Y^{T_0}\|_F^2}{\|Y^{T_0}\|_2^2} \leq \bar{r}_1$ $\frac{\|R^{T_0}\|_F^2}{\|R^{T_0}\|_2^2} \leq \bar{r}_2$, and their largest singular values are bounded above such that $\sigma_1(Y^{T_0}) \leq \bar{\sigma}_1$ and $\sigma_1(R^{T_0}) \leq \bar{\sigma}_2$, where $\bar{r}_1$, $\bar{r}_2$, $\bar{\sigma}_1$ and $\bar{\sigma}_2$ may depend on J and T_0.

The intuition behind Assumption 4 is better seen by considering the rank of the $J \times T_0$ design matrix Y^{T_0}. If the matrix had fixed rank $r < \min\{T_0, J\}$ all points would lie in a low dimensional manifold of the space and the pre-treatment fit error would grow proportional to r. Given that in our setting the error terms are $i.i.d$ shocks, this is not a reasonable assumption. Instead, we consider a bound on the numerical rank; the ratio between the Frobenius and 2-norm of a matrix. This notion of rank allows for points to lie "close" to a low dimensional manifold. Furthermore, for any matrix A it follows that

$$\frac{\|A\|_F^2}{\|A\|_2^2} \leq \mathrm{rank}(A),$$

therefore the bounded numerical rank assumption is implied by a bounded rank assumption. Whether Assumption 4 is satisfied will depend on the model primitives. In particular, it will be satisfied when the signal to noise ratio is high. That is, when the factor structure $\mu_i' F_t$ dominates the noise term ϵ. In cases in which σ_ϵ is large relative to the factor term the numerical rank is likely to be large and the pre-treatment fit bad. In section 6 we explore the performance of our estimator in a variety of settings and propose checks researchers can implement to evaluate whether their empirical setting is likely to satisfy this assumption.

Theorem 2 (Factor term consistency)**.** *Under Assumptions 1-4, for $t > T_0$ conditional on the unobserved components $\mu_i' F_t$ and A_{it}, the following bound holds for all J, T_1 and T_0*

$$\frac{1}{JT_1}\mathbb{E}\left[\left|\sum_{it} \tilde{Z}_{it}\tilde{\mu}_i' F_t\right|\right] \leq \left(\frac{\bar{F}^2 k c_z c}{\underline{\xi}}\right)\left(2c\sqrt{\frac{J}{T_0}}\sigma_\epsilon + (\bar{r}_1\bar{\sigma}_1 + \theta\bar{r}_2\bar{\sigma}_2)\left[\frac{1}{\sqrt{JT_0}} + C^2\sqrt{\frac{J}{T_0}}\right]\right)$$

where $c = 1 + C$ *and all other terms are defined in the assumptions. Furthermore, as* $JT_1 \to \infty$ *and* $(\bar{r}_1\bar{\sigma}_1 + \theta\bar{r}_2\bar{\sigma}_2)\sqrt{\frac{J}{T_0}} \to 0$*,*

$$\frac{1}{JT_1}\sum_{it} \tilde{Z}_{it}\tilde{\mu}_i' F_t \xrightarrow{p} 0.$$

Theorem 2 shows that the bias due to the factor term is $o_p(1)$ as long as $(\bar{r}_1\bar{\sigma}_1 + \bar{r}_2\bar{\sigma}_2)\sqrt{\frac{J}{T_0}} \to 0$. For fixed J, this implies that we need $T_0, T_1 \to \infty$. The restrictions on the rank and $\bar{\sigma}_1$, $\bar{\sigma}_2$ are not uncommon in the matrix completion literature. Combining the consistency result with the additional assumptions on the instrument behavior we can show that the synthetic IV estimator is a consistent estimator of θ.

Theorem 3 (Consistency)**.** *Under Assumptions 1-4, as* $JT_1 \to \infty$ *and* $(\bar{r}_1\bar{\sigma}_1 + \theta\bar{r}_2\bar{\sigma}_2)\sqrt{\frac{J}{T_0}} \to 0$*,*

$$\tilde{\theta}^{TSLS} - \theta \xrightarrow{p} 0,$$
$$\tilde{\theta}_Z^{TSLS} - \theta \xrightarrow{p} 0,$$
$$\tilde{\theta}_{YR}^{TSLS} - \theta \xrightarrow{p} 0.$$

Theorem 3 states that both the synthetic IV estimator $\tilde{\theta}^{TSLS}$ and the synthetic IV estimators for which we do not debias the instrument $\tilde{\theta}_{YR}^{TSLS}$ and for which we only debias the instrument $\tilde{\theta}_Z^{TSLS}$, are consistent given our assumptions and the rate conditions of Theorem 2. Under our model and assumptions, the standard TSLS estimator will not be consistent in general as $\frac{1}{JT_1}\sum_{it} Z_{it}\mu_i' F_t$ may not converge in probability to zero. As discussed, however, the synthetic IV estimator is biased in finite samples and the finite sample bias will depend on the signal to noise ratio, the length of the pre and post treatment periods in relation to the number of units J and, through the first stage, the correlation between Z_{it} and $\mu_i' F_t$. It is important to note that while debiasing the instrument does not affect the consistency of the estimator it may improve the finite sample properties of the estimator. In the appendix, we show under additional assumptions, that debiasing the instrument can lead to a stronger first stage and lower correlation with the debiased unobserved term $\tilde{\mu}_i' F_t$, leading to better finite sample properties. We confirm this intuition in the simulation study by comparing $\tilde{\theta}^{TSLS}$ and $\tilde{\theta}_Z^{TSLS}$. Finally, under Assumptions 1-4 it is also possible to show that the synthetic IV estimator is asymptotically normal.

Theorem 4 (Asymptotic normality)**.** *Under Assumption 1-4, conditional on weights w and instruments Z_{it}, if $\sqrt{\frac{T_1}{T_0}}(1+J)(\bar{r}_1\bar{\sigma}_1+\theta\bar{r}_2\bar{\sigma}_2)\to 0$ and $\frac{1}{\sqrt{JT_1}}\max_i\sum_{j\neq i}|w_{ji}|\to 0$, then as $JT_1\to\infty$*

$$\frac{\sqrt{JT_1}(\tilde{\theta}^{TSLS}-\theta)}{v_{JT_1}}\xrightarrow{d}(\gamma Q)^{-1}\times N(0,1).$$

where $v_{JT_1}^2=\frac{1}{JT_1}\sum_{it}var(\tilde{Z}_{it}\tilde{\epsilon}_{it}\mid Z,w)=\frac{1}{JT_1}\sum_{i,t>T_0}\sigma_\epsilon^2\tilde{\alpha}_{it}^2$ and $\tilde{\alpha}_{it}=\tilde{Z}_{it}-\sum_{j\neq i}\tilde{Z}_{jt}w_{ji}$.Furthermore, given $w_i\in\mathcal{W}$, a sufficient condition for $\frac{1}{\sqrt{JT_1}}\max_i\sum_{j\neq i}|w_{ji}|\to 0$ is that $\frac{J}{T_1}\to 0$ as $JT_1\to\infty$.

Theorem 4 shows that the estimator converges to a normal random variable centered at the true parameter when normalized by the conditional variance v_{JT_1} which depends on the instruments Z and synthetic control weights w. The result allows us to construct standard asymptotic confidence intervals by using the sample counterparts. Let $\tilde{\sigma}_{TSLS}^2$ denote the variance of the synthetic TSLS estimator which is given by

$$\tilde{\sigma}_{TSLS}^2=\frac{JT_1\hat{v}_{JT_1}^2}{(\sum_{i,t>T_0}\tilde{Z}_{it>T_0}\tilde{R}_{it})^2}=\frac{\hat{\sigma}_\epsilon^2\|\tilde{\alpha}\|_2^2}{(\sum_{i,t>T_0}\tilde{Z}_{it}\tilde{R}_{it})^2},$$

where $\tilde{\sigma}_\epsilon^2$ can be estimated from the regression residuals, and the denominator and $\|\tilde{\alpha}\|_2^2=\sum_{it}\tilde{\alpha}_{it}^2$ can be computed directly from the data. After computing this quantity, standard $(1-\alpha)\%$ confidence intervals can be constructed such that

$$\theta\in\left[\tilde{\theta}_{TSLS}-z_{1-\alpha/2}\times\frac{\tilde{\sigma}_{TSLS}}{\sqrt{JT_1}},\tilde{\theta}_{TSLS}+z_{1-\alpha/2}\times\frac{\tilde{\sigma}_{TSLS}}{\sqrt{JT_1}}\right],\tag{10}$$

where $z_{1-\alpha/2}$ denotes the $(1-\alpha/2)$-quantile of the standard normal distribution.

The result in Theorem 4 requires an additional density condition with respect to the conditions for consistency in Theorem 2. The condition requires that for the sequence of synthetic control weights w, $\frac{1}{\sqrt{JT_1}}\max_i\sum_{j\neq i}|w_{ji}|\to 0$. Intuitively, it ensures that the weights are not concentrated on a few units such that, as JT_1 grows, the estimator does not depend on a few data points. While, in general, whether this condition is satisfied will depend on the model primitives, under the l_1-norm constraint in program (3), a sufficient condition is that $J/T_1\to 0$ as $JT_1\to\infty$. In finite samples, however, we can inspect directly whether our estimated weights $\hat{w}$ are dense or not. In the simulation exercise and empirical applications we show that in general the weights are dense, explaining the good behavior of the SIV estimator in shorter panels.

Variance comparison. A natural question is how does the variance of the SIV estimator compare with the variance of the standard TSLS. Which variance is larger will depend on the model primitives and the relationship between the unobserved confounder U and the instrument Z. The intuition is similar to that of the effect of adding additional covariates on the variance of the OLS estimator for a parameter of interest in a linear model. Adding additional covariates reduces the variance of the unobserved component, but this may come at the cost of variance inflation due to the correlation between the regressor of interest and the additional covariates.

In our setting, if the SIV estimator is successful in partialling out the unmeasured confounder U, and the variance of U is large relative to σ_ϵ^2 or the correlation between U and Z is sufficiently small, then the SIV estimator will exhibit a smaller variance than the standard TSLS. To see this, suppose that in addition to Assumptions 1-4, $U_{it} = \mu_i' F_t$ are i.i.d random variables and $var(\mu_i' F_t) = \sigma_u^2 > 0$. Then, it follows that the asymptotic variances of the TSLS and SIV estimators are given by $\sigma_{TSLS}^2 = (\gamma^2 Q_Z)^{-1}[\sigma_u^2 + \sigma_\epsilon^2]$ and $\sigma_{SIV}^2 = (\gamma^2 Q)^{-1}\sigma_\epsilon^2$. Therefore, the ratio of the variances is greater than one when

$$\frac{\sigma_{TSLS}^2}{\sigma_{SIV}^2} = \frac{\sigma_u^2 + \sigma_\epsilon^2}{\sigma_\epsilon^2}\frac{Q}{Q_Z} > 1 \iff \frac{\sigma_u^2 + \sigma_\epsilon^2}{\sigma_\epsilon^2} > \frac{Q_Z}{Q}. \tag{11}$$

The LHS in the last expression in (11) is the signal to noise ratio, which given that $\sigma_u^2, \sigma_\epsilon^2 > 0$, is strictly greater than one. The RHS depends on the correlation between Z and U. Recall that Q is the residual variation in Z after U has been projected out, therefore $Q_Z/Q \geq 1$, with $Q_Z/Q = 1$ when $corr(Z, U) = 0$ and $Q_Z/Q \to \infty$ when $corr(Z, U) \to 1$ and $Q \to 0$. It follows that when $corr(Z, U) = 0$, the SIV estimator will have strictly lower variance than the TSLS, and when $corr(Z, U) \to 1$ the SIV will have greater variance than the TSLS. Intuitively, when $corr(Z, U) = 0$ this is exactly the case in which controlling for additional covariates would lower the variance of the OLS estimator of θ when R is randomly assigned (e.g. in an RCT). As the correlation between Z and U increases which variance dominates will depend on the signal to noise ratio. We investigate this trade off in the simulation exercise and the empirical applications and find that often the SIV exhibits lower variance than the TSLS.

5. Extensions

5.1. Combining SIV with additional estimators

The set up described under Assumptions 1 and 2 highlights a trade off between using the instrument variation to address the endogeneity bias due to the correlation between ϵ and η and incurring an omitted variable bias due to the instrument's correlation with the unobserved term $\mu_i' F_t$. The synthetic IV estimator can address these biases asymptotically in regimes in which σ_ϵ is small relative to the variation in $\mu_i' F_t$ as we highlighted in section 4. However, when σ_ϵ is large the endogeneity concern becomes more important than the omitted variable bias and, therefore, we might be able to design an estimator that addresses this bias more directly. With this in mind, we consider an additional estimator that will perform better in cases in which the noise level is high, and propose an ensemble estimator as a 'doubly robust' alternative to the synthetic IV.

Suppose that the instrument also follows a factor structure, such that $Z_{it} = Z_i' G_t$ for factor loadings Z_i and common factors G_t. This is the case in shift share designs such as the Syrian refugee example or the China shock study. In such cases, a natural estimator robust to noise ϵ_{it} is one that computes the synthetic control weights after projecting the outcome variable in the instrument space. The intuition for this estimator is that the outcome Y_{it} is noisy due to the unobserved error ϵ, but given our partial instrument validity assumption 2, after projecting the outcome in the instrument space we partial out the noise.

The *projected* synthetic estimator can be computed similarly to the SIV estimator, with an additional step.

1. Project to instrument space: for $t \leq T_0$, let $Y_{zt} = Z(Z'Z)^{-1}Z'Y_t$, where $Z = (Z_1, \ldots, Z_J)'$ and $Y_t = (Y_{1t}, \ldots, Y_{Jt})'$ are $J \times 1$ vectors.

2. Use the de-noised outcomes to compute the SC weights in the pre-period

$$\hat{w}_j^P \in \operatorname{argmin}_{w \in \mathcal{W}} \| Y_j^{T_0} - Y_{z,-j}^{T_0}{}' w \|^2.$$

3. Define the de-biased quantities $\tilde{Y}_{it}^P$, $\tilde{Z}_{it}^P$, $\tilde{R}_{it}^P$ accordingly for the projected weights $\hat{w}_j^P$.

4. For the post treatment period $t > T_0$, estimate the synthetic TSLS projected estimator

$$\tilde{\theta}^P = \left(\sum_{it} \tilde{Z}_{it}^P \tilde{R}_{it}^P \right)^{-1} \sum_{it} \tilde{Z}_{it}^P \tilde{Y}_{it}^P.$$

The performance of the projected estimator vis-a-vis the SIV will depend on the data generating process primitives. In cases in which the noise error term ϵ_{it} is more important than the factor term $\mu_i' F_t$ the projected estimator will perform favorably. On the other hand, if the factor structure (the signal) dominates, the projected estimator will perform worse than the synthetic IV as it will fit the factor structure $\mu_i' F_t$ worse. To see this, consider the bound in Theorem 1 for the projected estimator. In the appendix we show that with high probability under Assumptions 1-4

$$\frac{1}{JT_1}\mathbb{E}\left[\left|\sum_{i,t>T_0} \tilde{Z}_{it}^P \tilde{\mu}_i^{P\prime} F_t\right|\right] \leq \left(\frac{\bar{F}^2 k c_z c}{\xi}\right)\left(2c\sqrt{\frac{1}{T_0}}\sigma_\epsilon \right.$$
$$\left. + \mathbb{E}\left[\frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{Y}_{jt}^P|\right] + \theta\mathbb{E}\left[\frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{R}_{jt}^P|\right]\right),$$

where all the terms are defined in Assumption 3 and $c = 1 + C$. This bound differs from the bound in Theorem 1 in two important ways. First, the contribution of the noise term ϵ_{it}^2 to the overfitting bias changes from scaling with $\sqrt{\frac{J}{T_0}}$ to scaling with $\frac{1}{\sqrt{T_0}}$. This is because the estimated weights $\hat{w}_j^P$ dependence on the idiosyncratic shock ϵ_{it} vanishes asymptotically given that we are projecting into the instrument space and $Z_i \perp \epsilon_{it}$. Second, the pre-treatment fit will be worse as the weights only use variation in the instrument space. Therefore, while the projected estimator will be consistent under the assumptions and asymptotic regime of Theorem 2, its finite sample properties will differ from those of the SIV estimator. In particular, we expect the projected estimator to perform better than the SIV when

$$0 \leq \Delta_{T_0}^P - \Delta_{T_0}^{SIV} \leq \sqrt{J}\left(\frac{2(1+C)\sigma_\epsilon}{\sqrt{T_0}}\right), \tag{12}$$

where for an estimator $a \in \{\text{SIV}, \text{P}\}$ the expected pre-treatment fit is given by $\Delta_{T_0}^a = \mathbb{E}\left[\frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{Y}_{jt}^a|\right] + \theta\mathbb{E}\left[\frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{R}_{jt}^a|\right]$. Which means that in settings in which the noise level is large (high σ_ϵ) or the relative of number of units to pre-treatment periods is not close to zero, the projected estimator may have smaller finite sample bias than the SIV. While the bias itself cannot be evaluated directly in practice, by comparing differences in pre-treatment fit (LHS in 12) and checking the noise level of a given empirical setting (to evaluate the RHS in 12), we can have an idea of which estimator might be more or less biased.

Given that each estimator has a different finite sample bound, that depends on different model primitives, we can construct an ensemble estimator that combines both estimators and is robust to different sources of bias. For a hyper-parameter $\alpha^h \in [0,1]$ we define the ensemble estimator as

$$\tilde{\theta}^E(\alpha^h) = \alpha^h \tilde{\theta}^{TSLS} + (1-\alpha^h)\tilde{\theta}^P,$$

The α hyper-parameter can be chosen through cross-validation in the pre-period to optimize the mean squared error of the synthetic control estimator. The following steps detail how to compute the ensemble estimator.

1. Split the pre-period into a training period $1,\ldots,T_v$ and a validation period $T_v + 1,\ldots,T_0$.

2. In the training period compute the synthetic control weights for each estimator, $\hat{w}^P$ and $\hat{w}$, and the debiased outcomes $\tilde{Y}_{it}^P$ and $\tilde{Y}_{it}$.

3. In the validation period choose α^* to minimize the mean squared error in the validation

$$\frac{1}{J(T_0 - T_v)} \|\alpha^h \tilde{Y}^{P,T_v} + (1-\alpha^h)\tilde{Y}^{T_v}\|_2^2,$$

   where $\tilde{Y}^{T_v}$ denotes the debiased outcomes for the validation period.

4. Compute the ensemble estimator in the post period as $\alpha^* \tilde{\theta}^{TSLS} + (1-\alpha^*)\tilde{\theta}^P$.

In the appendix, we show that $\tilde{\theta}^E(\alpha^h)$ is a consistent estimator of θ for any $\alpha^h \in [0,1]$, and therefore, the cross-validated ensemble estimator is also consistent. The finite sample improvement of the cross-validated estimator will, however, depend on the finite sample bias differences between the estimators and the length of the validation period used to calibrate the hyper-parameter. In the simulation exercise in section 6, we show that the projected estimator performs well, with slightly worse performance to the SIV in low noise cases and better performance in high noise cases. In the Appendix (section 1.8) we also propose an alternative estimator based on a time-series aggregation scheme.

5.2. Alternative inference procedures

The asymptotic results in section 4 require that $JT_1 \to \infty$ and $J/T_0 \to 0$. However, in many shift-share IV and synthetic control design settings the researchers may have at their disposal a moderate number of units and time periods. This is the case of our main empirical application to the Syrian civil war. With this in mind, the literature has considered permutation based tests (Abadie et al. (2010), Abadie and Zhao (2022), Firpo and Possebom (2018)) and randomization inference procedures (Imbens and Rosenbaum (2005), Borusyak and Hull (2020)) as alternatives to asymptotic based confidence intervals. In this section, we describe how a split conformal inference procedure can be applied in the context of the synthetic IV. In the appendix, we describe an alternative randomization inference procedure.

Split conformal inference Our discussion in section 2 high-lights the use of reduce-form event studies as a way to assess if there is unmeasured confounding in an instrument $Z_{it} = Z_i \times \bar{H}_t$. In the Appendix (section 1.7) we expand our main theoretical set up to include the event study designs. The event studies take the following form

$$Y_{it} = \sum_{k \neq T_0} \theta_k (\mathbb{1}\{t = k\} \times Z_i) + \epsilon_{it}, \tag{13}$$

where, because $Z_{it} = 0$ for $t \leq T_0$, in absence of unmeasured confounding, we have that $\theta_l = 0$ for $l \leq T_0$. Therefore, in our linear IV design (1), testing the null $H_0 : \theta = 0$ is implied by testing that $\{\theta_l = \theta_k \text{ for all } l \leq T_0 \text{ and } k > T_0\}$. We propose a permutation based test for this null following the split-conformal inference procedures of Abadie and Zhao (2022) and Chernozhukov et al. (2021). The test can be implemented with any estimator of the event study coefficients θ_k, but we detail the procedure for the SIV estimator.

1. Split $1, \ldots, T_0$ into a *training* period $1, \ldots, T_b$ and a *blank* period $T_b + 1, \ldots, T_0$.
2. Compute SC weights in the training period and define debiased quantities accordingly.
3. Run reduced form event regression as in (13) using the debiased quantities $\tilde{Y}_{it}$ and $\tilde{Z}_i$ to get estimates $\{\tilde{\theta}_{T_b+1}, \ldots, \tilde{\theta}_T\}$.
4. Generate $T_1 \times 1$ permutation vectors $\theta_\pi = (\tilde{\theta}_{\pi(1)}, \ldots, \tilde{\theta}_{\pi(T_1)})$ for $\pi \in \Pi$, where Π is the set of size T_1 combinations from $T_b, \ldots, T$.
5. Compute the permutation test statistic $S(\theta) = 1/(T - T_0)\|\theta\|_1$ for each $\pi \in \Pi$.
6. Compute the permutation p-value:

$$\hat{p} = \frac{1}{\Pi} \sum_{\pi \in |\Pi|} \mathbf{1}(S(\tilde{\theta}_\pi) \geq S(\tilde{\theta}_{t > T_0}))$$

Under our design Assumptions 1-3, if additionally, $Z_{it} = Z_i \times \bar{H}_t$ and $\{\lambda_t\}_{t>T_b}$ is a sequence of exchangeable random variables independent of ϵ_{it} and η_{it}, it follows that tests based on $\hat{p}$ are exact in the sense that under the null H_0, for $\alpha \in [0,1]$ we have that

$$\alpha - \frac{1}{|\Pi|} \leq P(\hat{p} \leq \alpha) \leq \alpha$$

where P is taken over the distribution of $\{\epsilon_{it}, \lambda_t, \eta_{it}\}$. Note that given that Assumption 3 ensures that ϵ_{it} and η_{it} are $i.i.d$, it follows that for all i, $\{\epsilon_{it}, \lambda_t, \eta_{it}\}$ are exchangeable in t and the test exact validity result follows directly from Chernozhukov et al. (2021). For a result that relaxes the exchangeability assumption we refer readers to Abadie and Zhao (2022).

The permutation based inference procedure is going to be exactly valid in settings in which the time series structure of the unobserved confounder satisfies the exchangeability restriction. This is in contrast to the CI based on the SIV variance estimator (10) which are valid asymptotically under the regime of Theorem 4. In general, however, the power of the permutation based tests may be smaller and will depend on the number of time periods available and the noise levels. In the simulations, in section 6, we highlight the complementarity of both inference procedures in detecting the true effects using the SIV estimator.

6. Simulation study

In this section, we consider a simulation design calibrated to the Syrian empirical application. In the appendix, we consider different simulation designs with varying number of units, time periods, instrument strength and signal-to-noise ratios. All simulation designs follow the following data generating process

$$\begin{aligned}
Y_{it} &= \theta R_{it} + \mu_i' f_t + \epsilon_{it}, \\
R_{it} &= (\gamma Z_{it} + \eta_{it}) * \mathbb{1}(t \geq T_0), \\
Z_{it} &= Z_i' g_t * \mathbb{1}(t \geq T_0),
\end{aligned}$$

with time series structure

$$\begin{aligned}
f_t &= \kappa_f f_{t-1} + u_{ft}, \\
g_t &= \kappa_g g_{t-1} + u_{gt},
\end{aligned}$$

and error structure

$$\begin{pmatrix} u_{ft} \\ g_{ft} \end{pmatrix} \sim N\left(0, \begin{bmatrix} \sigma_f^2 & \rho_g \sigma_f \sigma_g \\ \rho_g \sigma_f \sigma_g & \sigma_g^2 \end{bmatrix}\right),$$
$$\begin{pmatrix} Z_i \\ \mu_i \end{pmatrix} \sim N\left(0, \begin{bmatrix} \sigma_z^2 & \rho_z \sigma_z \sigma_\mu \\ \rho_z \sigma_z \sigma_\mu & \sigma_\mu^2 \end{bmatrix}\right),$$
$$\begin{pmatrix} \epsilon_{it} \\ \eta_{it} \end{pmatrix} \sim N\left(0, \begin{bmatrix} \sigma_\epsilon^2 & \rho \sigma_\epsilon \sigma_\lambda \\ \rho \sigma_\epsilon \sigma_\lambda & \sigma_\lambda^2 \end{bmatrix}\right).$$

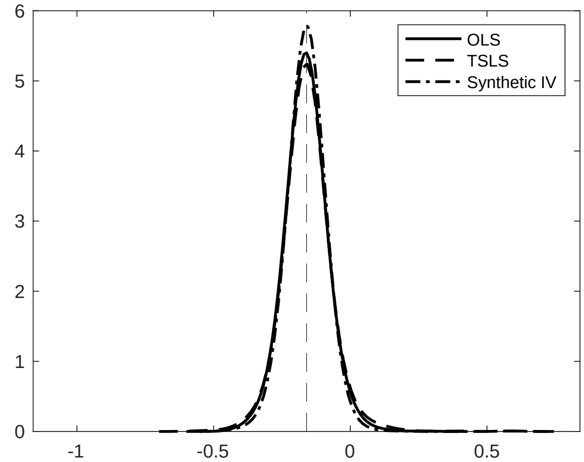
To map our simulation study to the data, we set the number of time periods to $T = 16$, with the intervention at $T_0 = 10$ and consider $J = 26$ regions. We target a relatively small true parameter of -0.16, with $\sigma_\epsilon^2 = \sigma_\eta^2 = 0.035$ calibrated to the residual variance in the data (noting that for some outcomes the variance is significantly smaller) and set σ_Z^2 and γ such that the F-statistic is 150. Finally, we let the signal be given by $\sigma_\mu^2 = 0.25$, consider one factor $k = 1$ (as it explains 98% of the variance in Y from PCA) and let the AR parameter be $\kappa = \kappa_f = \kappa_g = 0.5$. Given this design, we proceed by varying ρ, ρ_z, ρ_g, which given σ_Z^2 changes the correlation structure between the unobserved terms and the instrument (Q/Q_Z) and σ_ϵ, which given σ_μ changes the signal-to-noise ratio $\left(\frac{\sigma_\mu^2+\sigma_\epsilon^2}{\sigma_\epsilon^2}\right)$ to evaluate the performance of the estimators in different settings.

Figure 5 shows that the synthetic IV estimator is able to correct the bias present in the OLS and TSLS (with two-way fixed effects) estimators when there is endogeneity and omitted variable bias. Panel (a) shows the case in which all estimators are consistent (no correlation between R_{it} and U_{it} or ϵ_{it}). In this base case, as expected, the synthetic IV estimator performs similar to the OLS and TSLS estimators. In panel (b) we increase the correlation between ϵ_{it} and η_{it}, creating an endogeneity problem that can be addressed using the instrument. The OLS estimator is now biased, while the TSLS and synthetic IV estimators remain unbiased.[7] In panel (c) we introduce correlation between the instrument and the unobserved factor structure, by setting $\rho = \rho_z = \rho_g = 0.5$, and the instrument becomes invalid leading to biased TSLS estimates despite adding two-way fixed effects in the specification. The synthetic IV on the other hand is approximately unbiased, and we can reject that the estimated effect is zero at the 5% significance level. When we increase the correlation to $\rho = \rho_z = \rho_g = 0.7$ in panel (d), the bias in the OLS and TSLS estimators

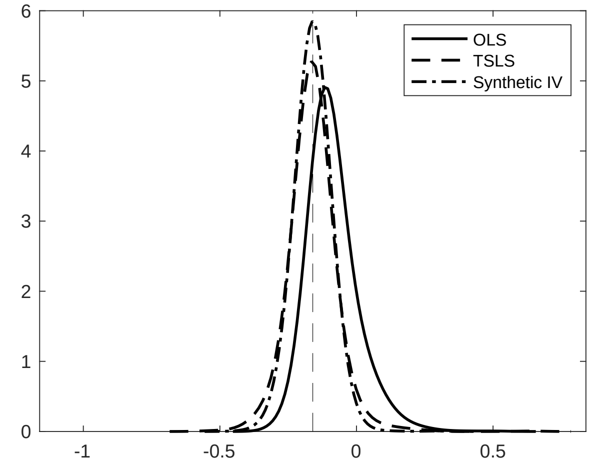
[7] In this design the bias due to the correlation between ϵ_{it} and η_{it} is small given that their variances are small relative to σ_μ and σ_z. In the simulation table and in the appendix we consider designs in which this bias is more important.

Figure 5: Model comparison in simulations

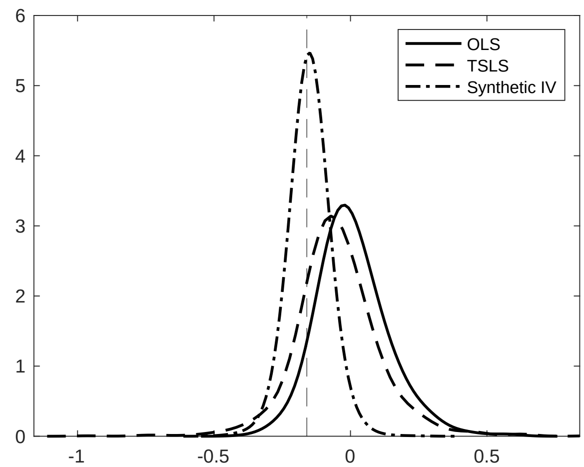
Note: Panels (a)-(d) display kernel density plots for TWFE OLS, TWFE TSLS and the synthetic IV. Panel (e) shows simulated event study estimates with 95% confidence bands for $\rho = \rho_z = \rho_g = 0.5$. *Panel (f) shows a histogram of* $\max_i \|w^i\|_1$ *and* $\sum_i \|w_i\|_2^2$ *for* $\rho = \rho_z = \rho_g = 0.5$. *Simulations are done over 10000 iterations with the parameters calibrated to the Syrian example.*



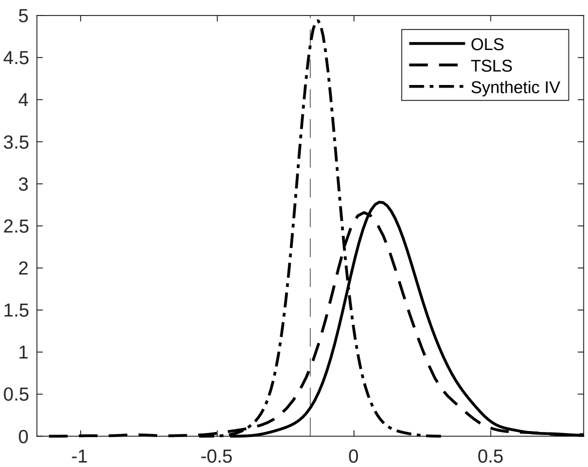
(a) $\rho = \rho_z = \rho_g = 0$



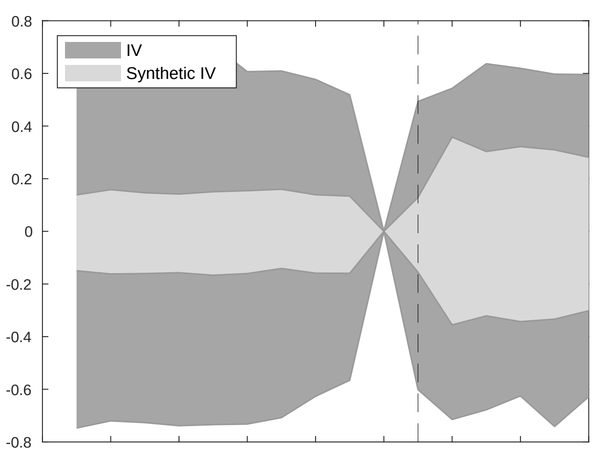
(b) $\rho = 0.5$, $\rho_z = \rho_g = 0$



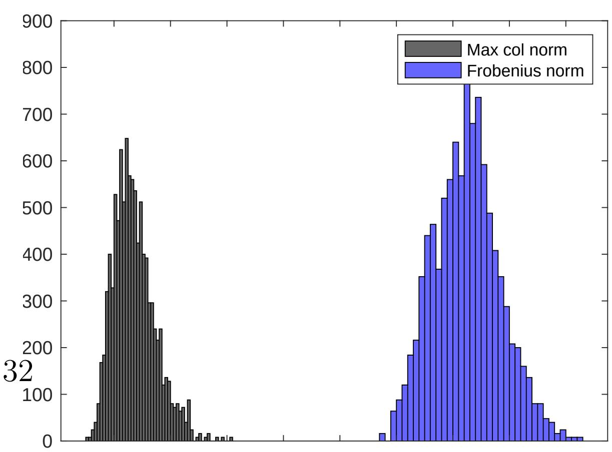
(c) $\rho = \rho_z = \rho_g = 0.5$



(d) $\rho = \rho_z = \rho_g = 0.7$



(e) Event study estimates



(f) Weight density

increases, but the synthetic IV continues to exhibit close to zero bias.

To relate this simulation results to the 'pre-trends' discussion in Figure 5, panel (e) shows the event study coefficients (over 10000 simulations). Before the treatment starts at $T_0 = 10$, the coefficients should be close to zero, as the instrument is not active. After the treatment starts the variation in the event study coefficients should increase given that $|\theta| > 0$.[8] The TSLS estimator shows large deviations in the pre period similar to those in the post period, indicating the presence of 'pre-trends' and unmeasured confounding. On the other hand, the synthetic IV limits the deviations in the pre period and also reduces the post period variation, suggesting that it is partialling out part of the unmeasured confounding. Finally, panel (f) in Figure 5 shows the density of the estimated synthetic control weights when $\rho = \rho_z = \rho_g = 0.5$. The max col norm histogram shows the value of $\max_i \|w^i\|_1 = \sum_{j\neq i} |w_{ji}|$ across simulations. It shows that the weights are dense, with no one unit receiving all weight across synthetic controls, and that the weight condition in Theorem 4 is likely to be satisfied as $\frac{1}{\sqrt{JT_1}} \max_i \sum_{j\neq i} |w_{ji}| \simeq 2.5/\sqrt{26 \times 6} \simeq 0.2$.

To gain further insight into the behavior of the different estimators, Table 1 shows the mean, variance, bias and mean-squared error of each estimator over 10000 simulations for different correlations and noise levels. In all cases, the SIV outperforms the OLS and TSLS (with two-way fixed effects) in terms of bias and mean-squared error, often by an order of magnitude. Furthermore, the SIV exhibits close to zero bias in settings with moderate noise and correlation levels. We compare the SIV to the synthetic IV for which we only debias the instrument $\tilde{\theta}_Z$ (SIV Z in the table), the projected SIV and the ensemble estimator proposed in section 5. As expected from the theoretical discussion, while $\tilde{\theta}_Z$ performs similar to SIV for moderate correlation settings, as the correlation grows (Q becomes smaller) the finite sample behavior of the SIV is better and $\tilde{\theta}_Z$ exhibits more bias. More so, also as expected, the projected SIV performs worse than the SIV in low noise settings and high correlation levels, but is robust to increasing the noise level. Intuitively, the projected SIV estimates (and mean TSLS estimates) do not change much as σ_ϵ is increased as the noise is orthogonal to the instrument. With this in mind, the ensemble estimator that combines the SIV and the projected SIV achieves better bias and MSE than the SIV in most settings.

Table 2 considers inference for the SIV estimator in our calibrated simulation design. The upper panel checks the coverage under the null $\theta = 0$ in the setting in which the

[8]Note that in this design across simulations we expect that the event study coefficients are centered at zero, which is why both IV and SIV are centered at zero. However, within a simulation the coefficients should be zero in the pre-period and non-zero in the post-period. Hence, the variation across simulations gives speaks to the bias in the coefficients in the pre and post periods.

Table 1: Simulations calibrated to the Syrian example for different $\rho = \rho_z = \rho_g = r$ and σ_ϵ.

	r=0.5				r=0.7				r=0.9			
	Mean	Var	Bias	MSE	Mean	Var	Bias	MSE	Mean	Var	Bias	MSE
					$\sigma_\epsilon = 1/2\sigma_{Syria}$							
OLS (TWFE)	-0.018	0.018	0.142	0.038	0.096	0.024	0.256	0.090	0.244	0.023	0.404	0.186
TSLS (TWFE)	-0.049	0.023	0.111	0.035	0.058	0.031	0.218	0.078	0.201	0.028	0.361	0.158
SIV	-0.155	0.002	0.005	0.002	-0.142	0.003	0.018	0.003	-0.087	0.008	0.073	0.013
projected SIV	-0.187	0.009	-0.027	0.010	-0.173	0.008	-0.013	0.008	-0.113	0.012	0.047	0.014
SIV + projected	-0.156	0.002	0.004	0.002	-0.144	0.003	0.016	0.003	-0.092	0.008	0.068	0.012
SIV Z	-0.142	0.003	0.018	0.004	-0.108	0.006	0.052	0.008	0.003	0.010	0.163	0.037
					$\sigma_\epsilon = \sigma_{Syria}$							
OLS (TWFE)	0.005	0.017	0.165	0.044	0.126	0.022	0.286	0.104	0.277	0.021	0.437	0.212
TSLS (TWFE)	-0.049	0.023	0.111	0.036	0.058	0.032	0.218	0.079	0.200	0.028	0.360	0.158
SIV	-0.151	0.004	0.009	0.004	-0.132	0.005	0.028	0.006	-0.056	0.013	0.104	0.024
projected SIV	-0.188	0.011	-0.028	0.012	-0.169	0.011	-0.009	0.011	-0.092	0.018	0.068	0.023
SIV + projected	-0.154	0.004	0.006	0.004	-0.137	0.005	0.023	0.006	-0.067	0.013	0.093	0.021
SIV Z	-0.135	0.005	0.025	0.005	-0.092	0.007	0.068	0.012	0.037	0.011	0.197	0.050
					$\sigma_\epsilon = 2\sigma_{Syria}$							
OLS (TWFE)	0.041	0.015	0.201	0.056	0.170	0.021	0.330	0.130	0.327	0.020	0.487	0.257
TSLS (TWFE)	-0.049	0.025	0.111	0.038	0.057	0.033	0.217	0.080	0.199	0.030	0.359	0.159
SIV	-0.144	0.008	0.016	0.008	-0.114	0.010	0.046	0.012	-0.013	0.020	0.147	0.042
projected SIV	-0.189	0.017	-0.029	0.018	-0.163	0.018	-0.003	0.018	-0.064	0.035	0.096	0.044
SIV + projected	-0.151	0.007	0.009	0.007	-0.125	0.009	0.035	0.011	-0.031	0.022	0.129	0.039
SIV Z	-0.125	0.008	0.035	0.009	-0.070	0.011	0.090	0.019	0.075	0.013	0.235	0.068
					$\sigma_\epsilon = 4\sigma_{Syria}$							
OLS (TWFE)	0.090	0.014	0.250	0.077	0.231	0.019	0.391	0.171	0.395	0.019	0.555	0.327
TSLS (TWFE)	-0.050	0.029	0.110	0.041	0.056	0.037	0.216	0.083	0.198	0.032	0.358	0.160
SIV	-0.132	0.027	0.028	0.028	-0.090	0.019	0.070	0.023	0.037	0.030	0.197	0.069
projected SIV	-0.187	0.029	-0.027	0.029	-0.152	0.034	0.008	0.034	-0.047	0.566	0.113	0.579
SIV + projected	-0.144	0.022	0.016	0.022	-0.109	0.019	0.051	0.022	0.000	0.200	0.160	0.226
SIV Z	-0.112	0.014	0.048	0.017	-0.045	0.017	0.115	0.030	0.110	0.017	0.270	0.090
					$\sigma_\epsilon = 8\sigma_{Syria}$							
OLS (TWFE)	0.147	0.013	0.307	0.107	0.302	0.016	0.462	0.229	0.474	0.016	0.634	0.418
TSLS (TWFE)	-0.053	0.038	0.107	0.050	0.052	0.047	0.212	0.092	0.193	0.042	0.353	0.167
SIV	-0.124	0.031	0.036	0.032	-0.062	0.041	0.098	0.050	0.081	0.049	0.241	0.107
projected SIV	-0.181	0.060	-0.021	0.061	-0.155	0.513	0.005	0.513	0.036	2.163	0.196	2.201
SIV + projected	-0.141	0.029	0.019	0.030	-0.095	0.093	0.065	0.097	0.063	0.486	0.223	0.536
SIV Z	-0.100	0.025	0.060	0.028	-0.020	0.028	0.140	0.047	0.139	0.024	0.299	0.113

common factors are exchangeable ($\kappa = 0$). We average over 10000 simulations the coverage of the 95% confidence intervals constructed according to 10 for the SIV estimator using the estimated variance $\tilde{\sigma}^2_{TSLS}$, the permutation p-value ($\hat{p}$) detailed in section 5 (with $T_b = 8$) and the weight density $\max_i \|w^i\|_1$ from the condition in Theorem 4. Overall, despite the small number of time periods and units ($T_1 = 6$, $T_0 = 10$, $J = 6$), CI based on the variance estimator $\tilde{\sigma}^2_{TSLS}$ exhibits close to nominal size in settings with moderate noise and moderate correlation with the unobserved confounder. In cases with low noise level and low correlation, the SIV exhibits slight over-coverage (as $\tilde{\sigma}^2_{TSLS}$ is inflated by the $\tilde{\alpha}_{it}$ terms), and in cases with high noise level and correlation, under-coverage. The under-coverage and finite sample bias for high correlation cases is both due to the rate of convergence depending on the correlation and a weaker first stage due to removing variation from the instrument (smaller Q). The permutation p-value, by construction, has nominal size regardless of the correlation or noise level, as can be seen by the average $\hat{p}$ being 0.5 across simulation designs. On the other hand, once we check the power of the test to detect the small $\theta = -0.16$ effect in the lower panel of Table 2, we find that the standard CI exhibit good power (60-90%) except in high noise cases, while the permutation p-value is under-powered. This should not be surprising given the small number of time periods and small effect. Overall, the good performance of the standard CI for the SIV estimator, in contrast to the theoretical rates, can be attributed to the synthetic controls balancing as the noise level increases. This can be seen by observing that the weight condition (given by $\max_i \|w^i\|_1$) improves with σ_ϵ and remains bounded for all simulation designs.

Table 2: Size and power simulations for $\rho = \rho_z = \rho_g = r$ and $\sigma_\epsilon = \iota\sigma_{Syria}$.

$H_0 : \theta = 0$ ($\kappa = 0$)

	r=0.0			r=0.5			r=0.7		
	Coverage	$\max_i \|w^i\|_1$	$\hat{p}$	Coverage	$\max_i \|w^i\|_1$	$\hat{p}$	Coverage	$\max_i \|w^i\|_1$	$\hat{p}$
ι									
0.100	0.998	2.585	0.508	0.997	2.595	0.489	0.990	2.570	0.495
0.500	0.986	2.418	0.510	0.963	2.419	0.493	0.950	2.405	0.498
1.000	0.956	2.322	0.509	0.938	2.318	0.496	0.898	2.305	0.499
2.000	0.912	2.213	0.499	0.890	2.210	0.498	0.822	2.202	0.502
4.000	0.878	2.123	0.504	0.829	2.119	0.495	0.732	2.117	0.511
8.00	0.847	2.061	0.508	0.800	2.059	0.499	0.675	2.060	0.501

$H_1 : \theta = -0.16$ ($\kappa = 0.5$)

	r=0.0			r=0.5			r=0.7		
	Power	$\max_i \|w^i\|_1$	$\hat{p} < 0.05$	Power	$\max_i \|w^i\|_1$	$\hat{p} < 0.05$	Power	$\max_i \|w^i\|_1$	$\hat{p} < 0.05$
ι									
0.100	0.910	2.593	0.568	0.865	2.609	0.502	0.855	2.602	0.410
0.500	0.865	2.439	0.294	0.794	2.442	0.260	0.788	2.455	0.187
1.000	0.813	2.349	0.203	0.707	2.347	0.178	0.712	2.360	0.130
2.000	0.726	2.247	0.150	0.592	2.247	0.127	0.592	2.252	0.104
4.000	0.619	2.153	0.114	0.447	2.150	0.090	0.455	2.154	0.086
8.00	0.490	2.085	0.088	0.347	2.081	0.090	0.319	2.080	0.088

The key takeaways from the simulations are that the SIV performs well in cases in which the TSLS and OLS do not, but that the estimator may be biased when the signal-to-noise level is weak (high noise) or the correlation with the unmeasured confounder is very large (no first stage). Another aspect highlighted by the theory and by reviews of best practices for synthetic control estimators (Abadie and Vives-i-Bastida, 2022), is how the relative sizes of T_0, T_1 and J influence the behavior of the estimator. The consistency result requires requires that both JT_1 is large and $\sqrt{J/T_0}$ is small. In our baseline simulation we considered a setting with $JT_1 = 156$ and $\sqrt{J/T_0} = 2.6$, but with a strong instrument and signal-to-noise ratio. In the appendix, we consider alternative simulation designs with different number of time periods, units, and weaker instrument and signal-to-noise ratios. Overall, for the different designs the same conclusions are drawn and the SIV estimator consistently outperforms the TSLS and OLS estimators (with two-way fixed effects).

With the simulation results in mind, we propose four robustness checks that practitioners should implement when using synthetic IV or similar estimators:

1. **Checking your first stage**: the debiasing procedure leads to a weaker first stage as variation is removed from the instrument. In cases with strong correlation between the instrument and the confounder (small Q) if the synthetic IV estimator exhibits a weak first stage researchers should be worried about using an IV strategy and the synthetic estimator.

2. **Checking your pre-treatment fit**: if the debiased outcomes exhibits large deviations in the pre-treatment period or an event study design reveals pre-trends, it is likely that the synthetic estimator will be biased and the signal-to-noise level too large for the estimator to perform well in the researcher's sample.

3. **Back testing**: given that the finite sample bias depends on the expected pre-treatment fit, back testing the intervention an evaluating the fit of the estimator, with weights computed in a training period, on a blank period can reveal whether the good pre-treatment fit was due to over-fitting (high noise) or due to partialling out the confounder. Additionally, researchers may implement the proposed permutation p-value to test the sharp null that the pre and post treatment event study estimates come from the same distributions.

4. **Checking the weight density**: ensuring that the synthetic control weights are not disproportionately weighting a few units by looking at the distribution of weights can reveal whether the asymptotic normality approximation is likely to be good in the researcher's empirical setting.

In the following section we implement these robustness checks when re-evaluating the effect of the Syrian refugee crisis using the synthetic IV.

7. Empirical applications

7.1. Revisiting the Syrian refugee shock

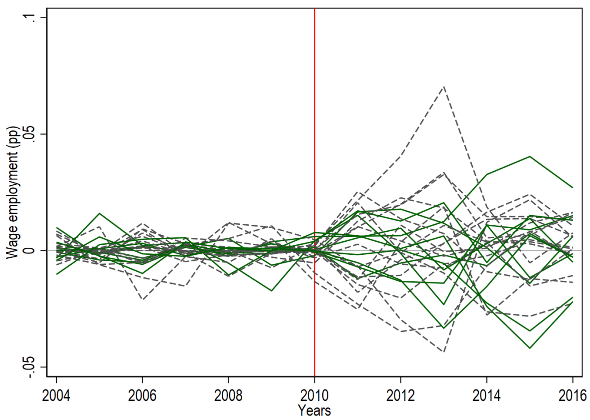
With the SIV tool at our disposal, we now re-visit our analysis of the impact of Syrian refugees on the salaried employment of low-skill natives. As detailed in section 3 we first solve the synthetic control problem (3) using the demeaned data between 2004–2010.[9] Then, we create synthetic regions with outcome Y^{SC}, treatment R^{SC}, and instrument Z^{SC} and debias the data by subtracting the raw data with the synthetic data, generating $\tilde{Y}_{it}$, $\tilde{R}_{it}$ and $\tilde{Z}_{it}$.

Before estimating the treatment effect via TSLS on the debiased data following step 2 in section 3, we implement the quality checks detailed in 6 to ensure that we are in a setting in which we can apply the SIV estimator. First, we check the matching quality in the pre-period, since as discussed in the theory section, goodness of fit is necessary to get consistent estimates using SIV. We plot the debiased wage-employment data ($\tilde{Y}$) in panel (a) Figure 6a, where black dashed lines belong to the less intensely treated regions that received less than 2% of refugees compared to their native population by 2016, and the green straight lines belong to the more intensely treated regions. During the training period 2004–2010, the debiased data is close to zero, which implies that we were able to match well on the pre-treatment trends.

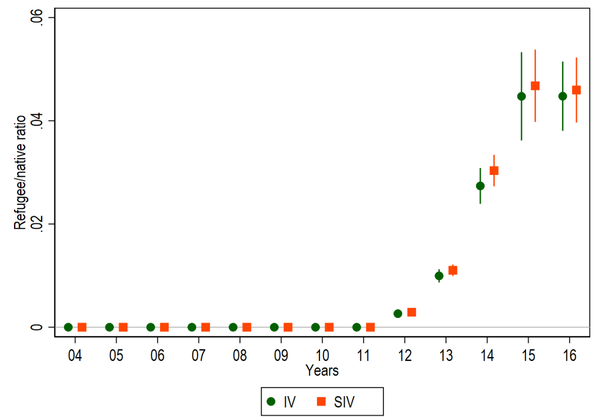
The second check we perform is to look at the first-stage using the debiased data, as the partialling out procedure will make the first stage weaker. We plot the first-stage estimates in panel (b) of Figure 6b. In our case, the debiased data maintains a strong first-stage. In a regression of $\tilde{R}$ on $\tilde{Z}$ while controlling for two-way fixed effects, the F-stat is 218. The third check we perform is to look at the reduced-form using the debiased data. If the matching was successful, i.e., the donor pool had regions with similar trends for all the regions in the sample, then the event-study design on the debiased data should find estimates around zero in the pre-period. As in the discussion of Figure 4, we estimate event-study regressions as in (9) for Z_i and $\tilde{Z}_i$ and plot the estimates Figure 6c. Adjusting for pre-rends, SIV finds

[9] Demeaning the individual regions is an important detail in the Turkish setting due to the large heterogeneity in development rates across regions. For example, Istanbul is the most developed region with the highest employment rate in Turkey. No convex combination of other regions can match Istanbul on levels, but matching on trends is feasible.

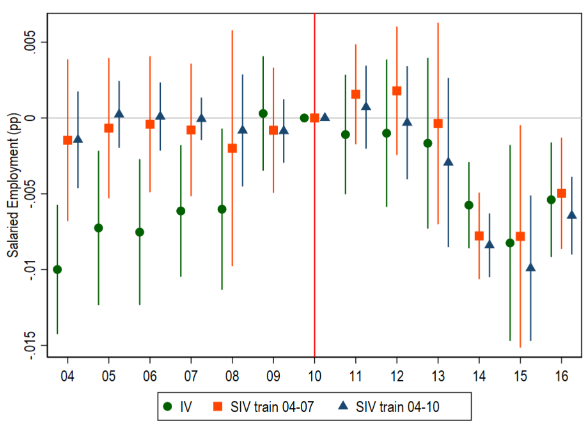
Figure 6: Quality Checks



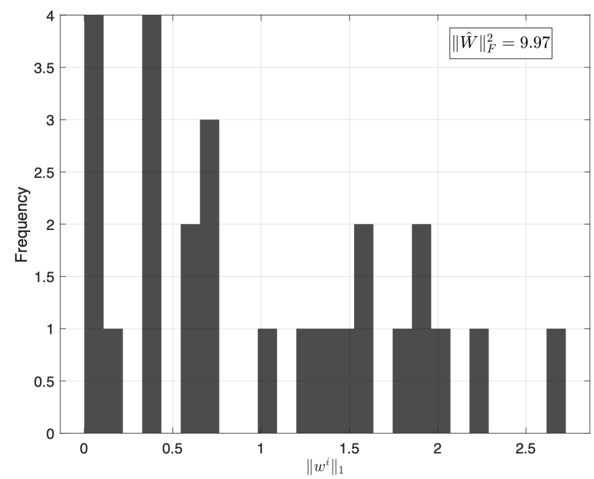
(a) Debiased Data



(b) Debiased First-Stage



(c) Debiased Reduced-Form



(d) Weight density

Notes: Panel (a) uses the debiased data. The green solid lines belong to the intensely treated regions, the black dashed lines belong to the rest, and the cutoff is 2% refugee/native ratio. The first-stage using both raw and debiased data is plotted in Panel (b). The F-stat in the main first-stage is 154 with the raw data and 218 with the debiased data. In Panel (c), the reduced-form estimates come from the event-study design shown in equation (9). The outcome variable is the wage-employment rate of low-skill natives. Pandel (d) shows $\|w^i\|_1$ for each Turkish region i, where w^i are the SC weights assigned to region i in the SC of the other regions. Standard errors are clustered at the region level. The 95% confidence interval is plotted.

slightly stronger disemployment effects in the post-period.

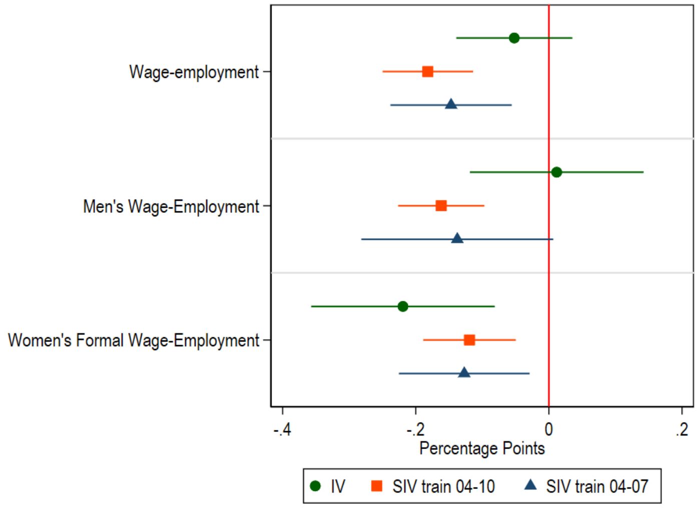
To test for over-fitting bias, we perform back-testing. In particular, instead of using the entire pre-period in the matching, we solve for the SC weights using data between 2004-2007 and follow the rest of the algorithm as specified before. We plot the estimates in Figure 6c in blue. Despite the reduced amount of time periods that we match on, the reduced form does not find any placebo effect in the pre-period. All the estimates between 2004–2010 are both quantitatively close to zero and statistically not significant, meaning synthetic distance is successfully capturing the unmeasured confounder in the pre-period. Furthermore, we compute the permutation p-value described in section 5 and find that $\hat{p} = 0.023$, meaning that we can reject at the 5% significance level that the post-treatment event study coefficients are zero and jointly equivalent to the pre-treatment coefficients.

Finally, Figure 6d shows the l_1 norm of the weights each region has across synthetic controls. Given that no one region receives a large amount of the weight, and the maximum value (2.6) is small relative to $\sqrt{JT_1} \approx 12.5$, we are confindent that the density condition of Theorem 4 is satisfied.

It is worth further discussing why IV and SIV estimates differ less in the post-period than in the pre-period in Figure 4 and Figure 6c. While it is impossible to know the exact nature of the unmeasured confounding, some likely explanations can help understand the nature of the pre-trend. As explained in Gulek (2023), the regions close to the Syrian border are less-developed than the rest. Between 2004–2010, Turkey's GDP per capita grew by 75%. The data suggests that the less developed southeast regions were "catching up" to the rest of Turkey with higher salaried employment growth rates. This aggregate growth period did not last as Turkey entered a recession in 2013. If economic growth in the pre-period was the main reason behind the pre-trends, it is likely that these pre-trends would not extrapolate into the post-period. The SIV is capturing this underlying change in the unobserved confounder by not changing the post estimates by a considerable margin. The presence of pre-trends does not necessarily imply a violation of the parallel trends assumption in the post-period. Although this principle is widely recognized in theoretical discourse, it often receives insufficient attention in empirical studies. This oversight may stem from the challenge of publishing research that identifies significant pre-trends without addressing them through parametric techniques, like controlling for linear trends. In our case, adjusting for linear trends would have caused us to overestimate immigrants' effect on natives' salaried jobs.

The degree by which SIV and IV estimates differ in the post depends on the persistence of

Figure 7: SIV and IV estimates.



Notes: This Figure plots the IV estimates ($\hat{\theta}_{TSLS}$) in green circles, the SIV estimates ($\tilde{\theta}$) in orange squares, and the backdated SIV estimates in blue triangles for the three main outcome measures. 95% confindence internvals are provided, in the case of SIV they are constructed using $\tilde{\sigma}_{TSLS}$. The permutation p-values ($\hat{p}$) are 0.023, 0.013 and 0.032 respectively.

the unobserved confounders. For different outcomes, IV and SIV estimates can differ more. To show this in our context, we estimate event studies for the immigrants' effect on the salaried employment of low-skill men and formal salaried employment of low-skill women. Figure A.2.3, in the appendix, plots the estimates. In the pre-period, whereas regions close to border are observing a relative increase in men's salaried employment as seen in panel (a), they observe a relative decrease in women's formal salaried employment as seen in panel (b). In both cases, SIV eliminates the pre-trends and adjusts the estimates in the post period.

Having seen how SIV addresses the pre-trend problem in the event-study designs and satisfies the recommended empirical checks, we continue by implementing the second step of the algorithm: we apply TSLS to the debiased data. We estimate the effect of Syrian refugees on low-skill natives' wage-employment, and show heterogeneity by sex and formality. We plot the estimates in Figure 7. As a benchmark, we first show the IV estimates. A researcher using IV would find no effect on natives' or men's salaried employment, and find negative effects on women's salaried employment. However, using SIV, we find that Syrian refugees lowered natives' salaried employment in all cases. A 1 pp increase in the refugee/native ratio decreases low-skill natives' salaried employment rate by 0.16 pp for men and 0.10 pp for women. As a robustness check, we also show the results that rely on estimated weights using the only 2004–2007 as a training period. The results remain quantitatively and qualitatively very similar. Accordingly, the permutation p-values reject the null effect ($\theta = 0$) for all three outcome measures at the 5% significance level.

It is worth highlighting how much our method impacts the economic conclusions in the empirical setting. Turkey hosts the largest number of refugees in the world. Turkey's three most treated exposed regions (the ones that received the most refugees) observed an increase in labor supply of more than 10% in just five years. Refugees, especially men, have a high propensity to work: 87% of prime-age men are "employed" in Turkey (Turkish Red Crescent and WFP, 2019). Despite this *large* labor supply shock, in a short enough time period where spatial markets are unlikely to equilibrate and despite male refugees' having higher employment rates than male natives, the standard IV finds no disemployment effects for native men. Theoretically justifying this result would require either completely flat labor demand curves (Borjas, 2003) or refugees to provide a substantial positive product demand shock (Borjas, 2014). There is very little empirical evidence for both, especially considering that Syrian refugees left most of their wealth behind while escaping a civil war. SIV reveals that this significant labor supply shock has caused native disemployment in the short run for both men and women, which is consistent with economic theory.

7.2. Revisiting the China Shock

SIV can be applied to any exposure and shift-share design. As an additional empirical example, we estimate the effect of Chinese imports on manufacturing employment in the United States following the identification strategy of Autor et al. (2013). The authors are interested in the following regression (where we omit covariates for simplicity)

$$\begin{aligned} Y_{it} &= \beta X_{it} + \epsilon_{it} \\ X_{it} &= \gamma Z_{it} + \eta_{it} \end{aligned} \tag{14}$$

where Y_{it} is the percentage point change in the manufacturing employment rate for region i in decade t, $X_{it} = \sum_k s_{ikt} g_{kt}^{US}$ is the import exposure, where s_{ikt} is the industry-location share at the beginning of period, and g_{kt} is a normalized measure of the growth of imports from China to the US in industry k. The import exposure to China is instrumented by the increase in Chinese imports by high-income countries: $Z_{it} = \sum_k s_{ikt-1} g_{kt}^{\text{high-income}}$, where s_{ikt-1} is the share of industry k in the previous period and $g_{kt}^{\text{high-income}}$ is a normalized measure of the growth of Chinese imports to selected high-income countries. We focus on the TSLS estimates from Tables 2 and 3 of Autor et al. (2013).[10]

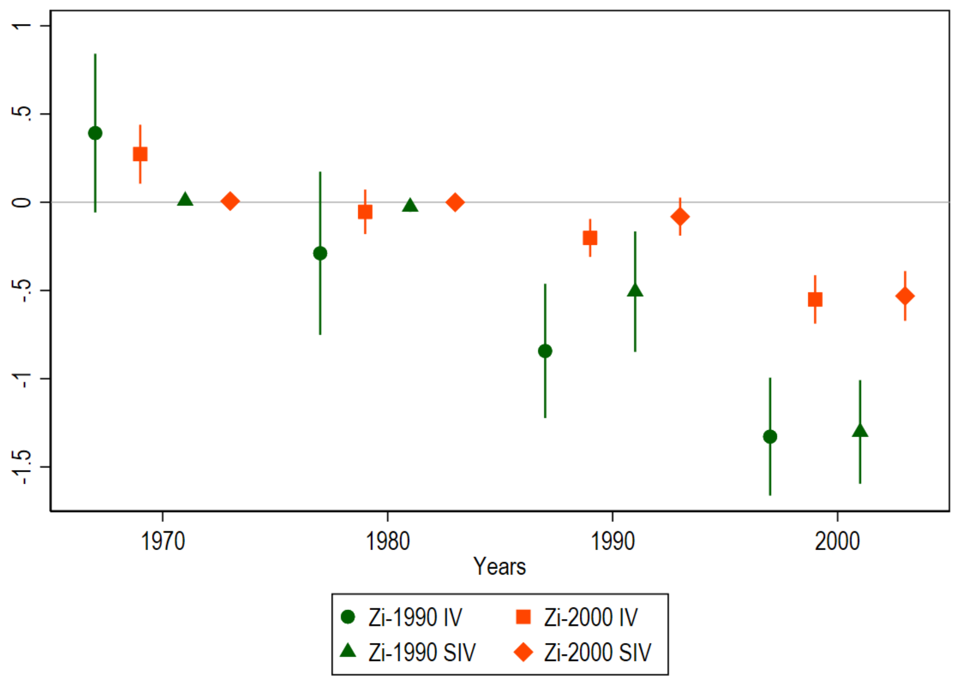
The paper considers 4 periods of data: 1970–1980, 1980–1990, 1990–2000, 2000-2007. We denote these periods by their starting year throughout the exercise (e.g., 1990 refers to the period between 1990–2000). The Chinese import shock takes place in 1990–2000 and 2000–2007. The growth in Chinese imports from high-income countries in 1990–2000, $g_{k,1990}^{\text{high-income}}$, predicts an exposure across US commuting zones via their pre-existing industry structure. We denote this exposure as $Z_{i,1990}$ and define $Z_{i,2000}$ similarly, as the exposure predicted by the growth in Chinese imports from high-income countries between 2000–2007. Together, $Z_{i,1990}$ and $Z_{i,2000}$ constitute the shift-share instrument used as $Z_{it} = Z_{i,1990}\mathbb{1}(t = 1990) + Z_{i,2000}\mathbb{1}(t = 2000)$. These two exposure measures have a correlation of 0.67 across 722 commuting zones, which implies that regions that have a high exposure in 1990 were also likely to have an high exposure in 2000. This suggests that the trade shocks in 1990 and 2000 were unlikely to be i.i.d., and hence we follow the exogeneity of shares assumption in this shift-share design (Goldsmith-Pinkham et al., 2020) as opposed to the exogeneity of shifts assumption in Borusyak et al. (2022).

Dissecting the shift-share instrument Z_{it} into its two "exposure" components $Z_{i,1990}$ and $Z_{i,2000}$ lends itself to an event-study design, as we already considered in section 2 (equation (2). Figure 3 shows that the exposure shares predict decreases in manufacturing employment in both 1990 and 2000, which is one of the core results in the China shock paper. However, this figure also reveals that the correlation between the exposure shares and manufacturing growth was positive in 1970, two decades before the Chinese shock, and has been decreasing since then. For example, the coefficient estimate of $Z_{i,1990}$ goes from 0.39 in 1970 to -0.28 in 1980. This pre-trend raises a concern regarding the validity of the *exogeneity of shares* assumption in this shift-share design because if this trend was to continue absent the China shock, we would have estimated the same "negative" employment effects in 1990 and 2000.

To apply the SIV estimator to the China shock example we follow the algorithm described in section 3. We first solve the synthetic control problem for all of the 722 CZs, where we match on the growth rate between 1970 and 1980, Then, we obtain the synthetic variables, y^{SC}, x^{SC}, Z^{SC}, and compute the debiased values $\tilde{Y}, \tilde{X}, \tilde{Z}$. Due to the small number of pre-periods and large number of donor units, the pre-treatment fit is almost perfect as can be seen in Figure 8 panel (b). As discussed in Abadie and Vives-i-Bastida (2022) in these settings it is likely that the synthetic control is fitting the noise, leading to over-fitting bias.

[10] In Table 2, the IV estimates without covariates and some placebo checks are shown. Table 3 shows the regressions with additional covariates.

Figure 8: Reduced-form estimates using the 1990 and 2000 shares



To address this in the appendix we re-do the analysis limiting the donor pool to the closest 100, 50, 30, and 20 donor regions according to the Euclidean distance.

We investigate the effects of the China shock on US manufacturing by comparing the IV and the SIV estimates. We find that the SIV results are slightly smaller in magnitude, but overall similar to the IV findings. Looking at 1990 in Figure 8 panel (b), we see that the SIV finds a statistically significant decrease in manufacturing employment due to Chinese imports, but the coefficient estimates are slightly smaller in magnitude than the IV estimates. In 2000, on the other hand, we find quantitatively the same results as ADH: adjusting for the pre-trends in 1970 and 1980 does not meaningfully change the estimates in 2000-2007.

Table 3 replicates the main findings in Autor et al. (2013). For the 1990–2000 period, the TSLS estimates suggest that a $1,000 increase in import exposure per worker leads to a decline in manufacturing employment of 0.89 pp. The SIV estimate for the same period is 33% smaller, but not statistically different, at 0.59 pp. This confirms the intuition from Figure 8 that adjusting for the pre-trend reduces slightly the China shock effect in 1990-2000. The results for the 2000-2007 period and the two periods combined (1990-2007) imply a decrease between 0.7 and 0.75 pp with little differences between TSLS and SIV. Autor et al. (2013) also report IV estimates with additional covariates. We replicate their results

Table 3: China shock effect

	1990–2000 (1)	2000-2007 (2)	1990–2007 (3)
2SLS	-0.888	-0.718	-0.746
	(0.181)	(0.064)	(0.068)
SIV	-0.588	-0.726	-0.703
	(0.198)	(0.070)	(0.067)

Notes: The first row replicates columns 1–3 of Table 2 in Autor et al. (2013). The second row presents the estimates using the SIV. The SC weights are estimated using the manufacturing growth rates in 1970 and 1980.

in the appendix in Table C.5 and find very comparable results between IV and SIV, revealing that the unmeasured confounding might be well proxied by the set of controls.

Overall, our replication implies that despite the strong pre-trend between 1970–1990, the IV estimates in 1990–2007 may not suffer from large biases due to unmeasured confounding. A potential explanation for this is that the decline in manufacturing employment growth suggested by the pre-trends flattens over time and disappears in the 2000s, leading to the small adjustment in the 1990-2000 period and no adjustment in the 2000-2007 period as reported by our SIV estimates.

7.3. The effect of search rankings

In this section we study an important question in platform and digital economics; what is the effect of product rankings on producer outcomes? Many digital platforms guide consumers to producers through search walls in which products are ranked according to ordered lists generated by a recommendation system. Some examples include search engines (e.g. Google, Bing), product market places (e.g. Amazon, Wayfair, Alibaba), food delivery platforms (e.g. Deliveroo, Uber Eats, Door Dash), travel comparison sites (e.g. Expedia, Google flights, Kayak) or social medial platforms (e.g. Facebook, Instagram, Tiktok). Given the importance of digital platforms, there is a growing literature in economics on the impact that rankings have on producer and consumer choices (e.g., Athey and Ellison (2011), Ursu (2018), Choi and Mela (2019), Hodgson and Lewis (2023), Compiani et al. (2024), Reimers and Waldfogel (2023)).

A key empirical challenge in the literature is that of estimating the causal effect of changing the rank of a producer from observational data. Given that ranking and recommendation systems are based on producer and consumer characteristics, observed ranks are endogenous. For example, top search results in Google search are determined through an auction, and therefore, advertisers with higher willingness to pay (which may be higher quality) are placed higher. Similarly, in food delivery platforms and social media platforms recommendation engines are trained on consumer and producer histories and therefore higher quality producers may receive higher rankings. This problem is well understood in the literature and there are a number of existing approaches to correct the endogeneity bias. For example, De los Santos and Koulayev (2017) consider a control function approach, Narayanan and Kalyanam (2015) and Moehring (2024) consider a regression discontinuity design, Ghose et al. (2014) specify and estimate a simultaneous equation model, and Rutz et al. (2012) consider a latent instrumental variables method. However, the validity of the different methods often relies on strong assumptions, or specific empirical settings in which the recommendation assignment function is known, which is why Ursu (2018) opts to analyze the results of a ranking A/B test in which ranks are fully randomized. In this paper, we propose a new set of instruments that can be used to study this question, show how the synthetic IV estimator can be applied to deal with unmeasured confounding and validate our observational results using an A/B test.

We provide new estimates of the effect of rank on producer sales in the empirical setting of a food delivery platform that selectively offers preferential search contracts to some producers. Producers with preferential search contracts see their position in the overall search wall of the platform fixed near the top (e.g. fixed at position 2,3,5,7,11) for a specific period of time. We use this type of *fixed* contracts as an instrument for rank. Given that many platforms offer preferential deals, sponsored slots, or promotion campaigns to producers, we believe that these kind of IVs might be available in many important digital settings. The fixed contract IV may be good for several reasons. First, given that the contracts directly change the ranks, the instrument is likely very relevant. Second, given that the contracts, in general, do not change other aspects of the producer's offering[11] the exclusion restriction may hold. However, it is unlikely that the producers offered the contracts where chosen at random, therefore the instrument may be invalid due to unmeasured confounding. This situation is common in IV approaches to ranking systems. Rutz et al. (2012) note that common

[11]In some cases commission fees are also simultaneously changed with the contract, but limited price pass-through is observed. See Vives-i-Bastida and Sabal (2024) for a model of the platform.

instruments used in these settings (e.g. lagged rankings or lagged outcomes) are unlikely to be valid due to omitted variable biases. In our setting, however, there is some hope. While the *fixed* contract IV may suffer from unmeasured confounding like the lagged IVs, we have at our disposal a pre-treatment period for each producer; the period before the producers were given a contract. Therefore, we are in a setting in which we may apply our proposed synthetic IV estimator.

Our data, provided by a large global delivery platform, consists of a sample of one million food delivery orders across six European cities between December 2022 and March 2023, involving 240 thousand customers and 1633 stores of which 183 received a fixed position contract (about 11%). The platform operates a marketplace for food delivery in which restaurants/stores are ranked in a search wall that customers browse before placing an order. The order/rank of the producers in the search wall depends on a recommendation system (unknown to the econometrician) involving producer specific characteristics and histories as well as order specific variables (such as geographical distance). Customers must scroll down to discover new producers, with only a few producers being shown at any given time on the screen (4 or 5 depending on the device used). The sample is chosen to include an A/B test the platform performed in which ranks were randomized at the customer level. In January and February 2023, 10936 customers where assigned to a treatment arm in which each week the rankings on their search wall were randomly generated within their market (city). We separate the A/B test data from the main dataset as a test set for validation. This allows us to compare different observational methods against a randomized experiment benchmark as in LaLonde (1986).

To study the effect of ranks on store outcomes, we aggregate the data at the store and week level. The aggregation can be micro-founded under a consumer discrete choice model in which consumer utility depends on common parameters that can be identified from aggregate market shares (Berry, 1994). For an in depth structural model of the platform, producer and consumer choices see Vives-i-Bastida and Sabal (2024) who use related data to study the welfare consequences of preferential contracts in platforms. The observed data consists of an outcome Y_{it}, the number of orders a producer i receives in week t, treatments R_{it}^k, the share of orders in which the producer was ranked in the top $k = 5, 8, 10$, and an instrument Z_{it} which consists of the share of orders in which the producer was *preferenced*. We do not consider producers with $Y_{it} = 0$ for some t, as we are interested in the effect of rankings for non-fringe producers that at least receive some orders on any given week. The instrument satisfies the requirements imposed by our design in Assumption 1, as $Z_{it} = 0$ for $t \leq T_0^i$, where T_0^i is the start of the contract period, but Z_{it} may vary both in time for $t > T_0^i$ and across i as producers have different preference contracts that vary across time. Therefore, this is an example of an instrument Z_{it} that does not have a factor structure as in the shift-share designs considered for the Syrian refugee and China shock studies. It is also important to note that in this context it is likely that SUTVA is violated, as receiving a higher rank implies other stores receive a lower rank, another reason to aggregate outcomes at the store-week level is to mitigate this concern. We note however, that the estimates of all of our estimators will suffer from this issue, making comparisons across estimators valid, and that we see our estimated effects as lower bounds on the true causal effect.

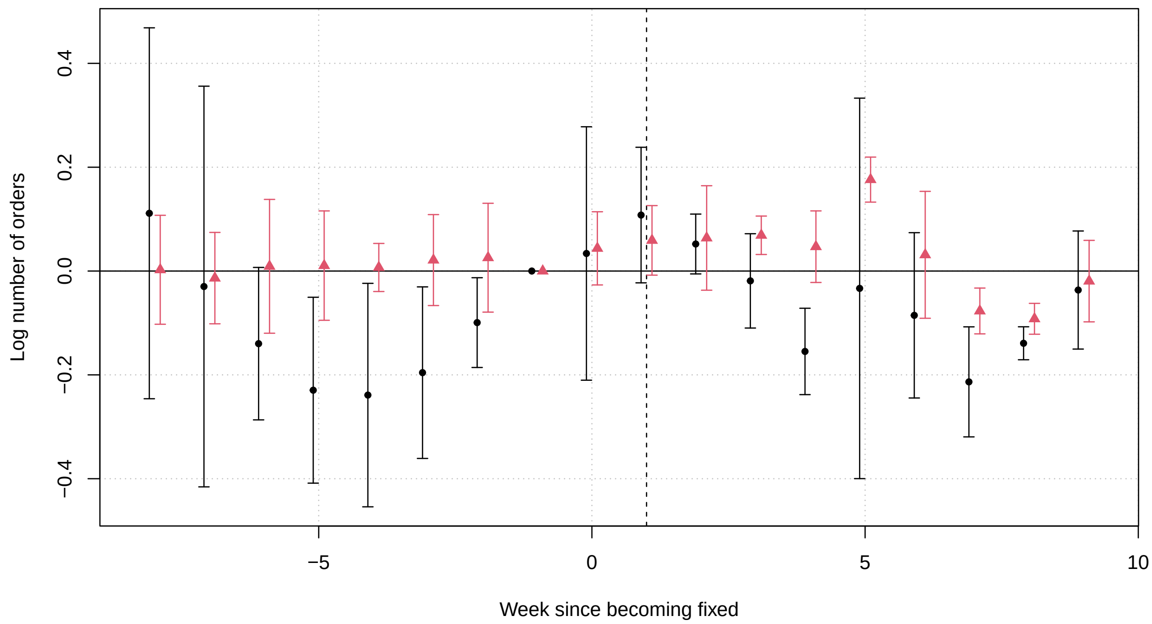
As in the other empirical examples we can assess the validity of the fixed contract instrument Z_{it} by checking the first stage strength and by inspecting the event study based on the reduced form. The first stage regression of R_{it}^k on Z_{it} (with two-way fixed effects) shows that, as expected, the instrument is a strong predictor of the rank of a store (see panel (a) of Figure A.4.5 in the appendix). Having a fixed contract increases the share of orders in which the store ranks in the top 5 by about 8%, in the top 8 by about 22% and in the top 10 by about 46%, with the first stage F-statistics being respectively, 158, 260 and 287. Therefore, we are confident that the fixed contract instrument is strong. To assess the potential unmeasured confounding in the reduced form, we estimate the following event study regression

$$log(Y_{it}) = \alpha_i + \delta_t + \sum_k \theta_k \mathbf{1}(t - T_0^i = k) + \epsilon_{it},$$

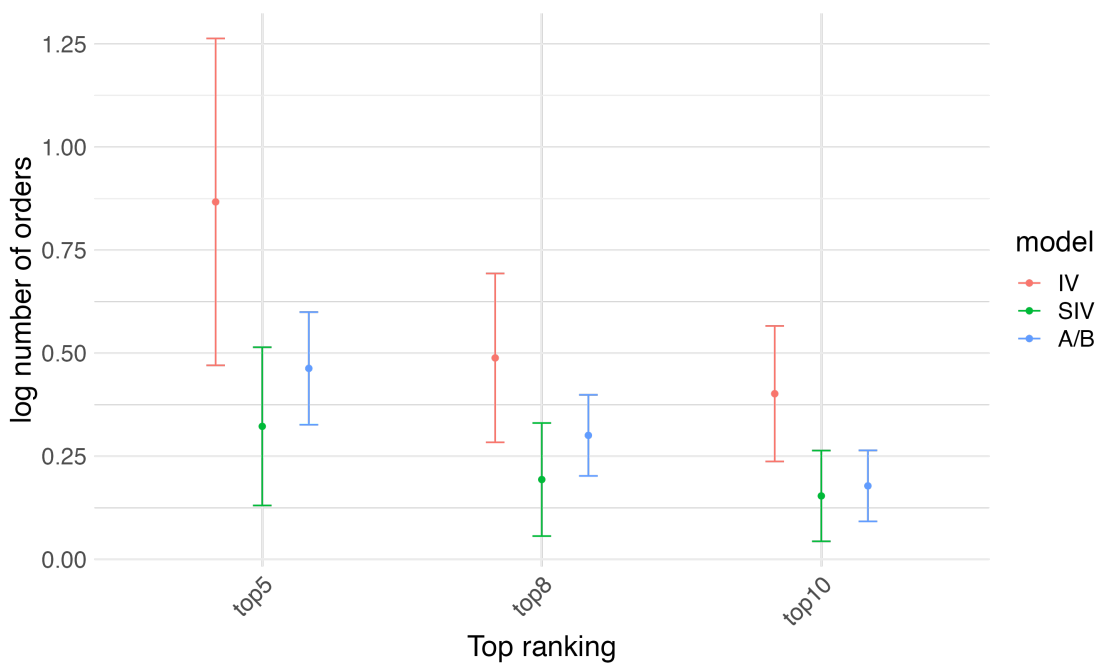
where T_0^i denotes the time of the start of the fixed contract by producer i. This regression does not use the variation across units and time of the *intensity* of the preferential contract, but it highlights the potential unmeasured confounders in determining which producers are preferenced and which are not. Given that adoption of fixed contracts is staggered, we use the imputation estimator of Sun and Abraham (2021) to account for potential dynamic effect contamination. Panel (a) of Figure 9 shows the estimated reduced form effects for the fixed contract IV in black. It can be appreciated that while becoming fixed seems to initially increase the number of orders by about 11% one period after treatment, there are clear non-linear (U-shaped) pre-trends, with similar patterns appearing in the post-treatment period. Intuitively, this may be due to the platform giving fixed contracts to stores that are in different trajectories, for example stores that are expected to increase their sales and are in an upward trend relative to others.

To deal for the unmeasured confounding highlighted by Figure 9 we apply the synthetic

Figure 9: Fixed contracts IV: reduced form and first stage



(a) Reduced form



(b) Rank effects estimates

Notes: Panel (a) shows the reduced form event studies for the fixed contract IV (black dots) and the synthetic IV (red triangles) using the Abraham and Sun (2020) estimator. The y-axis shows the number of orders in logs. Panel (b) shows the estimates of the effect of R_t^k for $k = 5, 8, 10$ on the log number of orders for the fixed contract IV, the synthetic IV and OLS for the A/B test sample. 95% confidence intervals are provided.

IV estimator. We modify the procedure described in section 3 to account for the staggered adoption of the fixed contracts. We compute the synthetic control weights for each fixed contract unit i using units that never receive a fixed contract and units that receive a fixed contract after unit i (i.e. units j such that $T_0^i \leq T_0^j$) as the donor pool. Then, we generate the debiased values $\tilde{Y}_{it}$, $\tilde{R}_{it}$ and $\tilde{Z}_{it}$ for $t > T_0^i$ and estimate the synthetic IV estimator for the post-treatment periods as in the standard case. Furthermore, given the large donor pool and number of treated units we both trim the donor pool to only match units close to each treated unit and drop the treated units with worst pre-treatment match. Our results, however, do not depend on this and are robust to different trimming operations. In panel (a) of Figure 9 we plot the synthetic IV event study estimates, using the debiased data, in red along with the standard IV estimates (in black). The synthetic IV does not exhibit pre-trends, as expected, but more importantly it also does not exhibit the U-shaped pattern in the post-treatment period that the standard IV does exhibit. The synthetic IV event study shows that after a producer becomes fixed the number of orders increases by about 10% over the following weeks.

Panel (b) of Figure 9 compares the fixed contract IV estimates, the synthetic IV estimates and the estimates from the A/B test. Given that in the A/B test, rank is randomized at the consumer level every week, the estimates should not suffer from omitted variable bias. We find that producer rank has a large effect on sales. The A/B test estimates (in blue) suggest that on average always being amongst the top 5 producers in the search wall on a given week leads to an increase in the number of orders of about 46% with respect to not being in the top 5. More so, as is common in the literature, we also find evidence of the effect decaying with the rank. Being in the top 8 leads to a smaller increase in sales of about 30%, and the effect for being in the top 10 is about 18%. When we compare the A/B test estimates to the fixed contract IV estimates, we see that there is likely a positive omitted variable bias despite the inclusion of two-way fixed effects. Respectively, the IV suggests a 86%, 49% and 40% increase in sales from being in the top 5, 8 and 10 ranks, more than double the A/B test estimates. This bias is in line with our reduced form discussion, since producers with higher growth potential may be more likely to receive a fixed contract. Reassuringly, the synthetic IV is successful in controlling the unmeasured confounding. The synthetic IV estimates (in green in Figure 9) are closer to the A/B test estimates, being respectively 33%, 18% and 14%. More importantly, a statistical test would not be able reject that the synthetic IV estimates are different from the A/B test for any of the rank treatments at a reasonable significance level. This is evidence that the proposed estimator appropriately controls the unmeasured confounding in a setting in which the IV estimator does not.

Overall, the application of the synthetic IV estimator to the rank effects example is relevant because (1) it shows how a new class of IVs can be used to study the question of how rankings affect producer outcomes in digital platforms (the fixed contract/promotion IVs) and (2) it shows that the synthetic IV estimator is effective in dealing with unmeasured confounding in a real setting in which we have access to an experiment to validate the observational estimates.

8. Conclusion

In this paper we provide a new method, the Synthetic IV, to deal with unmeasured confounding in panel data settings in which researchers have access to an instrumental variable that is only partially valid. By assuming a factor structure on the unobserved confounding term we derive conditions under which a synthetic IV estimator that combines Synthetic Controls and two-stage least squares is consistent and asymptotically normal. Through a simulation study, we show that the estimator performs well in a variety of empirical settings and removes the bias in cases in which TSLS and OLS with two-way fixed effects do not. We further showcase the applicability of SIV in three empirical examples: studying the effect of immigrants on labor markets using the Syrian refugee crisis in Turkey, studying the effect of Chinese imports on US manufacturing employment and studying the effects of rankings on producer outcomes in digital platforms.

References

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505.

Abadie, A., Diamond, A., and Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510.

Abadie, A. and Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review*, 93(1):113–132.

Abadie, A. and Imbens, G. W. (2012). A martingale representation for matching estimators. *Journal of the American Statistical Association*, 107(498):833–843.

Abadie, A. and Vives-i-Bastida, J. (2022). Synthetic controls in action.

Abadie, A. and Zhao, J. (2022). Synthetic controls for experimental design.

Adao, R., Kolesár, M., and Morales, E. (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics*, 134(4):1949–2010.

Agarwal, A., Dahleh, M., Shah, D., and Shen, D. (2021). Causal matrix completion.

Aksu, E., Erzan, R., and Kırdar, M. G. (2022). The impact of mass migration of syrians on the turkish labor market. *Labour Economics*, page 102183.

Anatolyev, S. and Mikusheva, A. (2022). Factor models with many assets: Strong factors, weak factors, and the two-pass procedure. *Journal of Econometrics*, 229(1):103–126.

Angrist, J. D. and Kugler, A. D. (2003). Protective or counter-productive? labour market institutions and the effect of immigration one unatives. *The Economic Journal*, 113(488):F302–F331.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., and Wager, S. (2021a). Synthetic difference-in-differences. *American Economic Review*, 111(12):4088–4118.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., and Wager, S. (2021b). Synthetic difference-in-differences. *American Economic Review*, 111(12):4088–4118.

Arkhangelsky, D. and Korovkin, V. (2023). On policy evaluation with aggregate time-series shocks.

Athey, S., Bayati, M., Doudchenko, N., Imbens, G., and Khosravi, K. (2021). Matrix completion methods for causal panel data models. *Journal of the American Statistical Association*, 116(536):1716–1730.

Athey, S. and Ellison, G. (2011). Position Auctions with Consumer Search*. *The Quarterly Journal of Economics*, 126(3):1213–1270.

Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The china syndrome: Local labor market effects of import competition in the united states. *American economic review*, 103(6):2121–2168.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279.

Bartel, A. P. (1989). Where do the new u.s. immigrants live? *Journal of Labor Economics*, 7(4):371–91.

Ben-Michael, E., Feller, A., and Rothstein, J. (2021). The augmented synthetic control method. *Journal of the American Statistical Association*. Forthcoming.

Berry, S. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics*, 25(2):242–262.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The quarterly journal of economics*, 118(4):1335–1374.

Borjas, G. J. (2014). Immigration economics. In *Immigration Economics*. Harvard University Press.

Borjas, G. J. (2017). The wage impact of the marielitos: A reappraisal. *ILR Review*, 70(5):1077–1110.

Borusyak, K. and Hull, P. (2020). Non-Random Exposure to Exogenous Shocks: Theory and Applications. NBER Working Papers 27845, National Bureau of Economic Research, Inc.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *The Review of Economic Studies*, 89(1):181–213.

Borusyak, K., Jaravel, X., and Spiess, J. (2023). Revisiting event study designs: Robust and efficient estimation.

Card, D. (1990). The impact of the mariel boatlift on the miami labor market. *ILR Review*, 43(2):245–257.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64.

Card, D. and Krueger, A. B. (2000). Minimum wages and employment: A case study of the fast-food industry in new jersey and pennsylvania: Reply. *The American Economic Review*, 90(5):1397–1420.

Cengiz, D. and Tekgüç, H. (2022). Is it merely a labor supply shock? impacts of syrian migrants on local economies in turkey. *ILR Review*, 75(3):741–768.

Chernozhukov, V., Wuthrich, K., and Zhu, Y. (2022). A t-test for synthetic controls.

Chernozhukov, V., Wüthrich, K., and Zhu, Y. (2021). An exact and robust conformal inference method for counterfactual and synthetic controls. *Journal of the American Statistical Association*. Forthcoming.

Choi, H. and Mela, C. F. (2019). Monetizing online marketplaces. *Marketing Science*, 38(6):948–972.

Compiani, G., Lewis, G., Peng, S., and Wang, P. (2024). Online search and optimal product rankings: An empirical framework. *Marketing Science*, 43(3):615–636.

Danieli, O., Nevo, D., Walk, I., Weinstein, B., and Zeltzer, D. (2024). Negative control falsification tests for instrumental variable designs. *Working Paper*.

de Chaisemartin, C., D'Haultfoeuille, X., Pasquier, F., Sow, D., and Vazquez-Bare, G. (2024). Difference-in-differences estimators for treatments continuously distributed at every period.

De los Santos, B. and Koulayev, S. (2017). Optimizing click-through in online rankings with endogenous search refinement. *Marketing Science*, 36(4):542–564.

Deaner, B. (2021). Proxy controls and panel data.

Doudchenko, N. and Imbens, G. W. (2016). Balancing, regression, difference-in-differences and synthetic control methods: A synthesis. NBER Working Paper, 22791.

Ferman, B. and Pinto, C. (2021). Synthetic controls with imperfect pre-treatment fit. *Quantitative Economics*. Forthcoming.

Firpo, S. and Possebom, V. (2018). Synthetic Control Method: Inference, Sensitivity Analysis and Confidence Sets. *Journal of Causal Inference*, 6(2):1–26.

Freyaldenhoven, S., Hansen, C., and Shapiro, J. M. (2019). Pre-event trends in the panel event-study design. *American Economic Review*, 109(9):3307–38.

Friedberg, R. M. (2001). The impact of mass migration on the israeli labor market. *The Quarterly Journal of Economics*, 116(4):1373–1408.

Ghose, A., Ipeirotis, P. G., and Li, B. (2014). Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science*, 60(7):1632–1654.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Gulek, A. (2023). Formal effects of informal labor supply: Evidence from the syrian refugees in turkey. Available at SSRN: https://ssrn.com/abstract=4264865.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1):25–46.

Hall, P. and Heyde, C. (1980). *Martingale limit theory and its application*. Probability and mathematical statistics. Academic Press.

Ham, D. W. and Miratrix, L. (2022). Benefits and costs of matching prior to a difference in differences analysis when parallel trends does not hold.

Hodgson, C. and Lewis, G. (2023). You can lead a horse to water: Spatial learning and path dependence in consumer search. Working Paper 31697, National Bureau of Economic Research.

Hunt, J. (1992). The impact of the 1962 repatriates from algeria on the french labor market. *ILR Review*, 45(3):556–572.

Imbens, G. and Rosenbaum, P. (2005). Randomization inference with an instrumental variable. *Journal of the Royal Statistical Society, Series A*, 168(1):109–126.

Imbens, G. W. and Viviano, D. (2023). Identification and inference for synthetic controls with confounding.

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018a). Shift-share instruments and the impact of immigration. Working Paper 24285, National Bureau of Economic Research.

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018b). Shift-share instruments and the impact of immigration. Technical report, National Bureau of Economic Research.

LaLonde, R. J. (1986). Evaluating the econometric evaluations of training programs with experimental data. *The American Economic Review*, 76(4):604–620.

Lebow, J. (2022). The labor market effects of venezuelan migration to colombia: reconciling conflicting results. *IZA Journal of Development and Migration*, 13(1).

Llull, J. (2017). Immigration, Wages, and Education: A Labour Market Equilibrium Structural Model. *The Review of Economic Studies*, 85(3):1852–1896.

Miao, W., Geng, Z., and Tchetgen Tchetgen, E. (2018). Identifying causal effects with proxy variables of an unmeasured confounder. *Biometrika*.

Moehring, A. (2024). Personalized Rankings and User Engagement: An Empirical Evaluation of the Reddit News Feed. OSF Preprints 8yuwe, Center for Open Science.

Mogstad, M. and Torgovitsky, A. (2024). Instrumental variables with unobserved heterogeneity in treatment effects. Working Paper 32927, National Bureau of Economic Research.

Narayanan, S. and Kalyanam, K. (2015). Position effects in search advertising and their moderators: A regression discontinuity approach. *Marketing Science*, 34(3):388–407.

Peri, G. and Yasenov, V. (2019). The labor market effects of a refugee wave synthetic control method meets the mariel boatlift. *Journal of Human Resources*, 54(2):267–309.

Reimers, I. and Waldfogel, J. (2023). A framework for detection, measurement, and welfare analysis of platform bias. Working Paper 31766, National Bureau of Economic Research.

Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*, 4(3):305–22.

Rudelson, M. and Vershynin, R. (2007). Sampling from large matrices: An approach through geometric functional analysis. *J. ACM*, 54(4):21–es.

Rutz, O. J., Bucklin, R. E., and Sonnier, G. P. (2012). A latent instrumental variables approach to modeling keyword conversion in paid search advertising. *Journal of Marketing Research*, 49(3):306–319.

Stephens Jr, M. and Yang, D.-Y. (2014). Compulsory education and the benefits of schooling. *American Economic Review*, 104(6):1777–1792.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199. Themed Issue: Treatment Effect 1.

Turkish Red Crescent and WFP (2019). Refugees in turkey: Livelihoods survey findings. https://reliefweb.int/report/turkey/refugees-turkey-livelihoods-survey-findings-2019-entr.

UNHCR (2021). Unhcr refugee statistics. https://www.unhcr.org/refugee-statistics/. Accessed: 2022-06-15.

Ursu, R. M. (2018). The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science*, 37(4):530–552.

Vives-i-Bastida, J. (2022). Predictor selection for synthetic controls.

Vives-i-Bastida, J. and Sabal, A. (2024). The effects of regulating food delivery platform design. Working project.

Wolfers, J. (2006). Did unilateral divorce laws raise divorce rates? a reconciliation and new results. *American Economic Review*, 96(5):1802–1820.

A.1. Theory

Throughout the appendix we introduce the following notation to refer to the debiased quantities: $\tilde{\epsilon}_{it} \equiv \epsilon_{it} - \sum_{j\neq i} \hat{w}_{ij}^{SC}\epsilon_{jt}$, as well as dropping the 'SC' weight subscript for notational convenience. Furthermore, we use T to mean T_1. The appendix consists of the following sections:

1. Bound on $\sum_{it} \tilde{Z}_{it}\tilde{\epsilon}_{it}$.
2. Proof of Theorem 1.
3. Proof of Theorem 2.
4. Proof of Theorem 3.
5. First stage debiasing lemmas.
6. Proof of Theorem 4.
7. Results for event study designs.
8. Results for projected and ensemble estimators.
9. Randomization inference.
10. Additional simulation tables.
11. Data appendix and additional tables for empirical examples.

1.1. Bound on $\sum_{it} \tilde{Z}_{it}\tilde{\epsilon}_{it}$

Lemma A.1 (Bound on $\tilde{Z}_{it}\tilde{\epsilon}_{it}$)**.** *Under Assumptions 1, 2, and 3, for any $\delta > 0$,*

$$\mathbb{P}\left(\left|\sum_{i,t>T_0} \tilde{Z}_{it}\tilde{\epsilon}_{it}\right| \geq \delta\right) \lesssim 2\exp\left(-\frac{\delta^2}{2c_z^2 JT\sigma_\epsilon^2}\right).$$

Hence, as $JT \to \infty$, $\frac{1}{JT}\sum_{it} \tilde{Z}_{it}\tilde{\epsilon}_{it} \xrightarrow{p} 0$.

Proof. First we show that the term has zero expectation given Assumption 3 and the independence of the error terms ϵ_{it}. The argument follows by noting that the SC weights depend only on ϵ_{it} for $t \leq T_0$,

$$w_j^{SC} \in \text{argmin}_{w\in\mathcal{W}} \|Y_j^{T_0} - Y_{-j}^{T_0}{}'w\|^2,$$

as only data from the pre-treatment period is used, here denoted as the $(J+1) \times T_0$ matrix Y^{T_0}. Therefore, $w_j^{SC} \perp \epsilon_{it}$ for $t > T_0$. Recall that by the law of iterated expectations if random variables b is independent of z and a such that $\mathbb{E}[b|c] = 0$ a.e., then $\mathbb{E}[ab|z] = 0$. Using this fact, under Assumption 2 it follows that $\mathbb{E}[\epsilon_{it} w_{ij}^{SC} | Z_{it}] = 0$ for $t > T_0$. Similarly, for any injective function $h : \text{Supp}(w) \to \mathbb{R}$ it follows that $h\left(w_j^{SC}\right) \perp \epsilon_{it}$ for $t > T_0$ and, consequently, $\mathbb{E}[\epsilon_{it} h\left(w_{ij}^{SC}\right) | Z_{it}] = 0$. To apply these facts, we re-write the second term, dropping the 'SC' subscript for convenience

$$
\begin{aligned}
\mathbb{E}\left[\tilde{Z}_{it}\left(\epsilon_{it} - \sum_{j \neq i} \hat{w}_{ij}^{SC} \epsilon_{jt}\right)\right] &= \mathbb{E}\left[(Z_{it} - Z_{-it}' w_i)(\epsilon_{it} - \epsilon_{-it}' w_i)\right] \\
&= \mathbb{E}\left[Z_{it} \mathbb{E}[\epsilon_{it} - \epsilon_{-it}' w_i | Z_{it}] - Z_{-it}' \mathbb{E}[w_i(\epsilon_{it} - \epsilon_{-it}' w_i) | Z_{it}]\right],
\end{aligned}
$$

where the $-i$ subscripts denote $J \times 1$ vectors not including unit i and w_i denote the $J \times 1$ vector of weights for unit i. Given that $\mathbb{E}[\epsilon_{it} w_{ij} | Z_{it}] = 0$ and $\mathbb{E}[\epsilon_{it} w_{ij}^2 | Z_{it}] = 0$, it follows that both conditional expectation terms are zero. Recall that for two vectors u and v, the following inequality holds $|u'v| \leq \|u\|_\infty \|v\|_1$. Therefore, Assumption 2, it follows that for $w \in \mathcal{W}$ for all i, t

$$
|Z_{it} - Z_{-it}' w_i| \leq \|Z_t\|_\infty (1 + \|w_i\|_1) \leq c_z(1 + C),
$$

where Z_t is the $J \times 1$ vector of instrument values at time t. It follows by the triangle inequality that

$$
\left|\sum_{it} \tilde{Z}_{it} \tilde{\epsilon}_{it}\right| \leq (1 + C) c_z \left|\sum_{it} \tilde{\epsilon}_{it}\right|.
$$

Given that ϵ_{it} are subgaussian and for $t > T_0$ independent of w_i, it follows that $\epsilon_{it} - \sum_{j \neq i} \hat{w}_{ij}^{SC} \epsilon_{jt}$ is a linear combination of subgaussian random variables. The first term has variance proxy σ_ϵ^2 and the second term has variance proxy $\sigma_\epsilon \|w_i\|^2 \leq C^2 \sigma^2$ as our weights satisfy $\|w_i\|^2 \leq \|w_i\|_1^2 \leq C^2$. Therefore, $\tilde{\epsilon}_{it}$ is subgaussian with variance proxy $2\sigma_\epsilon^2 C^2$. The result then follows directly by Hoeffding's inequality for subgaussian random variables (Theorem 2.6.2 Vershynin 2018)

$$
\begin{aligned}
\mathbb{P}\left(\left|\sum_{it} \tilde{Z}_{it} \tilde{\epsilon}_{it}\right| \geq \delta\right) &\leq \mathbb{P}\left(\left|\sum_{it} \tilde{\epsilon}_{it}\right| \geq \delta / ((1 + C) c_z)\right) \\
&\lesssim 2 \exp\left(-\frac{\delta^2}{2C^2 (1 + C)^2 c_z^2 J T \sigma_\epsilon^2}\right).
\end{aligned}
$$

□

1.2. Proof of Theorem 1

Proof. We start by re-writing the factor structure in terms of the outcome variable and in the pre-treatment period

$$\tilde{\mu}_i' F_t = \tilde{Y}_{it} - \theta \tilde{R}_{it} - \tilde{\epsilon}_{it}.$$

Using the projection trick, we can rewrite $\tilde{\mu}_i$ in terms of pre-treatment quantities:

$$\tilde{\mu}_i = (F_{T_0} F_{T_0}')^{-1} F_{T_0} (\tilde{Y}_i^{T_0} - \theta \tilde{R}_i^{T_0} - \tilde{\epsilon}_i^{T_0}).$$

With this in mind, consider the object of interest

$$\begin{aligned}
\left| \sum_{it} \tilde{Z}_{it} \tilde{\mu}_i' F_t \right| &= \left| \sum_{it} \tilde{Z}_{it} F_t' (F_{T_0} F_{T_0}')^{-1} F_{T_0} (\tilde{Y}_i^{T_0} - \theta \tilde{R}_i^{T_0} - \tilde{\epsilon}_i^{T_0}) \right| \\
&\leq \sum_{it} |\tilde{Z}_{it} F_t' (F_{T_0} F_{T_0}')^{-1} F_{T_0} \tilde{\epsilon}_i^{T_0}| + \sum_{it} |\tilde{Z}_{it} F_t' (F_{T_0} F_{T_0}')^{-1} F_{T_0} \tilde{Y}_i^{T_0}| + \theta \sum_{it} |\tilde{Z}_{it} F_t' (F_{T_0} F_{T_0}')^{-1} F_{T_0} \tilde{R}_i^{T_0}| \\
&\leq (1+C) c_z \left(\sum_{it} |F_t' (F_{T_0} F_{T_0}')^{-1} F_{T_0} \tilde{\epsilon}_i^{T_0}| + \sum_{it} |F_t' (F_{T_0} F_{T_0}')^{-1} F_{T_0} \tilde{Y}_i^{T_0}| + \theta \sum_{it} |F_t' (F_{T_0} F_{T_0}')^{-1} F_{T_0} \tilde{R}_i^{T_0}| \right).
\end{aligned}$$

Where the inequalities follow from the triangle inequality and the bound for $|\tilde{Z}_{it}| \leq (1+C) c_z$ which we derived in Lemma A.1. For the first term bound we proceed as in Abadie and Zhao (2022) and apply the Cauchy-Schwartz inequality and the eigenvalue bound on the Rayleigh quotient to bound the factor terms for any t, s

$$(F_t' (F_{T_0} F_{T_0}')^{-1} F_s)^2 \leq \left(\frac{\bar{F}^2 k}{T_0 \underset{\sim}{\xi}} \right)^2,$$

To bound these terms in expectation observe that $\bar{\epsilon}_{it} \equiv F_t' (F_{T_0} F_{T_0}')^{-1} F_{T_0} \epsilon_{iT_0}$ is a linear combination of subgaussian random variables and therefore it is itself a subgaussian random variable with variance proxy $\left(\frac{\bar{F}^2 k}{T_0 \underset{\sim}{\xi}} \right)^2 \frac{\sigma_\epsilon^2}{T_0}$, where for notational convenience we use $\epsilon_{iT_0} \equiv \epsilon_i^{T_0}$. Therefore,

$$\begin{aligned}
|\mathbb{E}[(\bar{\epsilon}_{iT_0} - \bar{\epsilon}_{-iT_0}' w_i)]| &\leq \mathbb{E}[|(\bar{\epsilon}_{iT_0} - \bar{\epsilon}_{-iT_0}' w_i)|] \\
&\leq (1+C) \mathbb{E}\left[\sum_j |\bar{\epsilon}_{iT_0}| \right] \\
&\leq (1+C) \left(\mathbb{E}\left[\sum_j |\bar{\epsilon}_{iT_0}|^2 \right] \right)^2 \\
&= (1+C) \left(\sum_j \mathbb{E}\left[|\bar{\epsilon}_{iT_0}|^2 \right] \right)^2
\end{aligned}$$

$$\leq 2(1+C)\left(\frac{\bar{F}^2 k}{\underline{\xi}}\right)\sqrt{\frac{J}{T_0}}\sigma_\epsilon.$$

The first inequality follows from Jensen's inequality. The second inequality follows by the triangle inequality, the absolute value and expectation operator inequality and $\|w_i\| \leq C$. The third follows from Holder's inequality with $q = 2$ and Jensen's inequality. Finally, the last inequality follows from Rigollet and Hutter 2019 (Lemma 1.4) which bounds absolute moments of sub-gaussian random variables. It follows that

$$\begin{aligned}\mathbb{E}[|F_t'(F_{T_0}F_{T_0}')^{-1}F_{T_0}(\epsilon_{iT_0} - \epsilon_{-iT_0}' w_i)|] &= \mathbb{E}[|(\bar{\epsilon}_{iT_0} - \bar{\epsilon}_{-iT_0}' w_i)|] \\ &\leq 2(1+C)\left(\frac{\bar{F}^2 k}{\underline{\xi}}\right)\sqrt{\frac{J}{T_0}}\sigma_\epsilon.\end{aligned}$$

For the term involving $\tilde{Y}$ we get that

$$\begin{aligned}\sum_{it} |F_t'(F_{T_0}F_{T_0}')^{-1}F_{T_0}\tilde{Y}_i^{T_0}| &\leq \left(\frac{\bar{F}^2 k}{\underline{\eta} T_0}\right)\sum_{it}|\sum_{t\leq T_0}\tilde{Y}_{it}| \\ &= \frac{T}{T_0}\left(\frac{\bar{F}^2 k}{\underline{\eta}}\right)\sum_{i,t\leq T_0}|\tilde{Y}_{it}|\end{aligned}$$

Dividing by TJ it follows that the term is bounded by

$$\left(\frac{\bar{F}^2 k}{\underline{\eta}}\right)\frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{Y}_{it}| = \left(\frac{\bar{F}^2 k}{\underline{\eta}}\right)MAD(\tilde{Y}^{T_0}).$$

where $MAD(\tilde{Y}^{T_0}) = \frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{Y}_{it}|$. An equivalent bound can be derived for the term involving $\tilde{R}$ which will depend on $\theta MAD(\tilde{R}^{T_0})$. The bound then follows from the proof of Theorem 1 and by Jensen's inequality applied to the absolute value. Consistency follows by an application of Markov's inequality. □

1.3. Proof of Theorem 2

Proof. The proof follows the proof of Theorem 1 by bounding $\mathbb{E}\left[\frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{Y}_{jt}|\right]$. It is useful to re-write the SC problem in matrix form. Let W be the $J \times J$ matrix of weights from program (3) where each row's sum is bounded by C and $diag(W) = 0$. Then the $J \times T_0$ matrix $\tilde{Y}^{T_0}$ can be re-written as $Y^{T_0} - \hat{W}Y^{T_0}$. It follows that the Frobenius norm of the matrix $\|\tilde{Y}^{T_0}\|_F^2 = \sum_{it<T_0}\tilde{Y}_{it}^2$ is bounded as follows

$$\|\tilde{Y}^{T_0}\|_F^2 = \|Y^{T_0} - \hat{W}Y^{T_0}\|_F^2 \leq \|Y^{T_0}\|_F^2 + \|\hat{W}Y^{T_0}\|_F^2$$

$$\begin{aligned}&\leq \|Y^{T_0}\|_F^2 + \|\hat{W}\|_F^2\|Y^{T_0}\|_F^2\\&\leq \|Y^{T_0}\|_F^2(1+C^2J)\\&\leq \bar{r}_1\bar{\sigma}_1(1+C^2J).\end{aligned}$$

where the first inequality follows by the triangle inequality. The second by the bound on the frobenius norm of a matrix product. The third by noting that for $w_i \in \mathcal{W}$ each row of W summed is bounded by C and $\|w_i\|^2 \leq \|w_i\|_1 \leq C^2$ which implies $\|\hat{W}\|_F^2 \leq JC^2$. Finally, the last inequality follows from A4 as $\|Y^{T_0}\|_F^2 \leq \|Y^{T_0}\|_2^2\bar{r}$. Next, observe that $\sum_{it<T_0}|Y_{it}| = \|vec(Y^{T_0})\|_1^2$ and $\|vec(Y^{T_0})\|_2^2 = \|Y^{T_0}\|_F^2$. So by the inequality between l_1 and l_2 norms,

$$\sum_{it\leq T_0}|Y_{it}| = \|vec(Y^{T_0})\|_1^2 \leq \sqrt{JT_0}\|vec(Y^{T_0})\|_2^2 = \sqrt{JT_0}\|Y^{T_0}\|_F^2.$$

Given the previous derivations we get the following bound

$$\frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{Y}_{jt}| \leq \frac{\sqrt{JT_0}}{JT_0}\bar{r}_1\bar{\sigma}_1(1+C^2J) = \bar{r}_1\bar{\sigma}_1\left(\frac{1}{\sqrt{JT_0}} + C^2\sqrt{\frac{J}{T_0}}\right).$$

A similar derivation for the term involving $\tilde{R}$ yields

$$\frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{R}_{jt}| \leq \frac{\sqrt{JT_0}}{JT_0}\bar{r}_2\bar{\sigma}_2(1+C^2J) = \bar{r}_2\bar{\sigma}_2\left(\frac{1}{\sqrt{JT_0}} + C^2\sqrt{\frac{J}{T_0}}\right).$$

The first part of the result then follows by noting that for a bounded random variable $|X| \leq K$, $\mathbb{E}|X| \leq K$. The consistency part follows by an application of Markov's inequality. □

1.4. Proof of Theorem 3

Lemma A.2 (U projection consistency). *Under the assumptions of Theorem 3 it follows that as $JT \to \infty$*

$$\frac{1}{JT}\|(I-\hat{W})U\|_1 \xrightarrow{p} 0,$$

where $\|\cdot\|_1$ denotes the sum of the absolute values elements of the matrix, I denotes the $J\times J$ identity matrix, $\hat{W}$ denotes the $J\times J$ matrix of weights from solving program 3 with $diag(\hat{W}) = 0$ and U denotes the $J\times T$ matrix of unobserved factors $\mu_i' F_t$. Furthermore, it follows that as $JT\to\infty$,

$$\frac{1}{JT}\|\hat{W} - P_U\|_F \xrightarrow{p} 0,$$

where $\|\cdot\|_F$ denotes the Frobenius norm, the sum of squared elements of the matrix, and P_U denotes the projection matrix $U(U'U)^{-1}U'$ for $J \times T$ matrix U.

Proof. From the proof of Theorem 1 we have that under the Assumptions 1-4 as $JT \to \infty$

$$\frac{1}{JT}\sum_{i,t>T_0} |\tilde{\mu}_i' F_t| \xrightarrow{p} 0.$$

Re-writing this statement in matrix form yields the first part of the proof. Let U be the $J \times T$ matrix with entries $\mu_i' F_t$ and $\hat{W}$ be the $J \times J$ matrix with zero diagonal elements and with each rows $\hat{w}_i^{*\prime}$, where $\hat{w}_i^*$ is the vector with entries equal to the weight vector $\hat{w}_i$ from program (3) with a zero in position i. Therefore, we have

$$\sum_{i,t>T_0} \tilde{\mu}_i' F_t = \sum_{i,t>T_0} [(I-\hat{W})U]_{it} = \|(I-\hat{W})U\|_1,$$

and given the results from Theorem 1 it follows that as $JT \to \infty$

$$\frac{1}{JT}\|(I-\hat{W})U\|_1 = \frac{1}{JT}\sum_{i,t>T_0} |\tilde{\mu}_i' F_t| \xrightarrow{p} 0,$$

which given that the LHS is non-negative and the RHS converges to zero in probability, implies the first part of the lemma.

For the second part, we are interested in the following object

$$\begin{aligned}\hat{W} - P_U &= \hat{W} - P_U + \hat{W}P_U - \hat{W}P_U \\ &= (I-P_U)\hat{W} - (I-\hat{W})P_U.\end{aligned}$$

By the triangle inequality we need to show that the following two objects on the LHS converge in probability to zero

$$\|\hat{W} - P_U\|_F \leq \|(I-\hat{W})P_U\|_F + \|(I-P_U)\hat{W}\|_F. \tag{A.1}$$

Define a linear operator $g_U : \mathbb{R}^{T\times J} \to \mathbb{R}^{J\times J}$ such that $g_U(A) := A(U'U)^{-1}U'$. Under Assumption 3 we have that $(U'U)^{-1}$ is invertible as the common factor matrix $F_T F_T'$ has bounded lowest eigenvalue and the factor loadings μ_i are bounded. Furthermore, given the bounds on the factor terms we have that g_U is a well defined bounded continuous linear operator. It follows by an application of the continuous mapping theorem for bounded functions that as $JT \to \infty$

$$\frac{1}{JT}\|(I-\hat{W})P_U\|_1 = \frac{1}{JT}\|g_U((I-\hat{W})U)\|_1 \xrightarrow{p} 0,$$

and convergence in the Frobenius norm follows from the l_1-l_2 inequality (for a vector a, $\|a\|_2 \leq \|a\|_1$). A similar argument can be applied to show that the second term in A.1 converges in probability, by noting that $P_U = P_U'$ and therefore that the Frobenius norm of both terms is the same. $\square$

Proof. Under the assumptions, Lemma A.1 and Theorem 3 show that both $\frac{1}{JT}\sum_{it}\tilde{Z}_{it}\tilde{\epsilon}_{it}$ and $\frac{1}{JT}\sum_{it}\tilde{Z}_{it}\tilde{\mu}_i' F_t$ are $o_p(1)$. It remains to be shown that the first stage term $\sum_{it}\tilde{R}_{it}\tilde{Z}_{it}$ does not go to zero in probability. Observe that

$$\begin{aligned}\frac{1}{JT}\sum_{it}\tilde{R}_{it}\tilde{Z}_{it} &= \frac{1}{JT}\sum_{it}(\gamma\tilde{Z}_{it}+\tilde{A}_{it}+\tilde{\eta}_{it})\tilde{Z}_{it}\\ &= \gamma\frac{1}{JT}\sum_{it}\tilde{Z}_{it}^2+\frac{1}{JT}\sum_{it}\tilde{Z}_{it}\tilde{A}_{it}+\frac{1}{JT}\sum_{it}\tilde{Z}_{it}\tilde{\eta}_{it}.\end{aligned}$$

Under Assumptions 1-3 we have that $\frac{1}{JT}\sum_{it}\tilde{Z}_{it}\tilde{\eta}_{it} = o_p(1)$ by Lemma A.1, given that the same assumptions for ϵ_{it} apply to η_{it}. Similarly, given $Z_{it} \perp A_{it}$, that Z_{it} is bounded, and $\frac{1}{JT}\sum_{i,t\geq T_0} A_{it} \xrightarrow{p} 0$ as $JT \to \infty$, it follows that $\frac{1}{JT}\sum_{it}\tilde{Z}_{it}\tilde{A}_{it} = o_p(1)$. Next, we show that the first term does not vanish in probability. Consider a $T \times J$ matrix Z for the instrument Z_{it} and an equivalent matrix $\tilde{Z}$ for the debiased instruments $\tilde{Z}_{it}$. We can express $\tilde{Z}$ in terms of Z and $\hat{W}$, the $J \times J$ matrix with zero diagonal elements and with each row $\hat{w}_i^{*\prime}$, where $\hat{w}_i^*$ is the vector with entries equal to the weight vector $\hat{w}_i$ from program (3) with a zero in position i,

$$\tilde{Z} = (I-\hat{W})Z = (I-\hat{W}+P_U-P_U)Z = (I-P_U)Z+(P_U-\hat{W})Z.$$

Recall that by the triangle inequality, for a norm $\|\cdot\|$ and matrices A, B it follows that $\|A+B\| \geq |\|A\|-\|B\||$. Using this inequality we have that

$$\begin{aligned}\|\tilde{Z}\|_F &= \|(I-P_U)Z+(P_U-\hat{W})Z\|_F\\ &\geq |\|(I-P_U)Z\|_F-\|(P_U-\hat{W})Z\|_F|\\ &\geq \|(I-P_U)Z\|_F-\|(P_U-\hat{W})Z\|_F\\ &\geq \|(I-P_U)Z\|_F-\|(P_U-\hat{W})\|_\infty\|Z\|_1.\end{aligned}$$

where the last inequality follows from the generalized Holder inequality for matrices, with the norms respectively representing maximum row and column sums. Given that $|Z_{it}| \leq c_z$ for all i,t it follows that $\|Z\|_1 \leq Jc_z$. By Lemma A.2 we have that as $JT \to \infty$

$$\frac{1}{JT}\|(P_U-\hat{W})\|_1 \xrightarrow{p} 0, \tag{A.2}$$

and given that P_U is symmetric it follows that the same is true for the infinity norm $\|\cdot\|_\infty$, an note that given Lemma A.2 it is sufficient to have that $T, T_0 \to \infty$ for fixed J. Therefore, we have that as $T, T_0 \to \infty$ $\frac{1}{JT}\|(P_U - \hat{W})\|_F^2 \|Z\|_F^2 \leq \frac{c_z}{T}\|(P_U - \hat{W})\|_1 \xrightarrow{p} 0$. For the term involving P_U, let $M_U = I - U_{JT_1}(U'_{JT_1}U_{JT_1})^{-1}U'_{JT_1}$ and $Z_{JT} = vec(Z)$, then

$$Z'_{JT}M_U Z_{JT} = Z'_{JT}M'_U M_U Z_{JT} = (M_U Z_{JT})'(M_U Z_{JT}) = \|M_U Z_{JT}\|_2^2 = \|(I - P_U)Z\|_F^2,$$

as M_U is idempotent. Combining (A.2) with $\frac{1}{JT}Z'_{JT}M_U Z_{JT} \xrightarrow{p} Q > 0$ as $JT \to \infty$ from Assumption 4, we have that as $JT \to \infty$

$$\frac{1}{JT}\sum_{it}\tilde{Z}_{it}^2 = \frac{1}{JT}\|\tilde{Z}\|_F^2 \geq Q - o_p(1).$$

It follows that this term is bounded below in probability. A similar argument yields that it is also bounded above in probability, since by the triangle inequality

$$\begin{aligned}\frac{1}{JT}\|\tilde{Z}\|_F^2 &= \frac{1}{JT}\|(I - P_U)Z + (P_U - \hat{W})Z\|_F^2 &\quad (A.3)\\ &\leq \frac{1}{JT}\||(I - P_U)Z\|_F^2 + \frac{1}{JT}\|(P_U - \hat{W})Z\|_F^2. &\quad (A.4)\end{aligned}$$

which converges in probability to Q as $JT \to \infty$ by the same argument as above. It follows by the dominated convergence theorem that $\frac{1}{JT}\|\tilde{Z}\|_F^2 \xrightarrow{p} Q$. Therefore, we have that the first stage terms are $O_p(1)$. Finally, consider the synthetic IV estimator decomposition

$$\begin{aligned}\tilde{\theta}^{TSLS} &= \theta + \left(\frac{1}{JT}\sum_{i,t>T_0}\tilde{Z}_{it}\tilde{R}_{it}\right)^{-1}\frac{1}{JT}\sum_{i,t>T_0}\tilde{Z}_{it}\left(\mu_i - \sum_{j\neq i}\hat{w}_{ij}^{SC}\mu_j\right)' F_t\\ &\quad + \left(\frac{1}{JT}\sum_{i,t>T_0}\tilde{Z}_{it}\tilde{R}_{it}\right)^{-1}\frac{1}{JT}\sum_{i,t>T_0}\tilde{Z}_{it}\left(\epsilon_{it} - \sum_{j\neq i}\hat{w}_{ij}^{SC}\epsilon_{jt}\right)\\ &= \theta + O_p(1)o_p(1) + O_p(1)o_p(1)\\ &= \theta + o_p(1)\end{aligned}$$

Given that in the proof of Theorem 2 we bound away the contribution of the instrument $\tilde{Z}_{it}$, the same proof can be applied for Z_{it} for $\tilde{\theta}_{YR}^{TSLS}$. Furthermore, observe that

$$\tilde{Z}'Z = Z'(I - W)'Z = Z'(I - P_U)Z + Z'(P_U - \hat{W})'Z \quad (A.5)$$

and that because P_U is symmetric and idempotent, a corollary of Lemma A.2 is that $\frac{1}{JT}\|\hat{W}' - P_U\|_1 \to 0$ in probability as $JT \to \infty$ since $\|\hat{W}' - P_U\|_F = \|(\hat{W} - P_U)'\|_F = \|\hat{W} - P_U\|_F$. Therefore, it follows that as $JT \to \infty$

$$\frac{1}{JT}\|(P_U - \hat{W})'Z\|_F^2 \xrightarrow{p} 0$$

by the same argument as for $\tilde{\theta}^{TSLS}$ and the same probability limit holds for the first stage term.

To show the consistency for $\tilde{\theta}_Z^{TSLS}$ observe that the equivalent decomposition (6) for the estimator involves bounding in probability the following term

$$\sum_{it} |\tilde{Z}_{it}\mu_i' F_t| = \sum_{it} |[Z'(I - \hat{W})'U]_{it}| = \|Z'(I - \hat{W})'U\|_1,$$

where $\|\cdot\|_1$ denotes the sum of the absolute value of the entries norm. Given that P_U is symmetric, a corollary of Lemma A.2 is that $\frac{1}{JT}\|(I - \hat{W})'U\|_1 \to 0$ in probability as $JT \to \infty$ since $\|\hat{W}' - P_U\|_F = \|(\hat{W} - P_U)'\|_F = \|\hat{W} - P_U\|_F$. Therefore, by bounding Z_{it}, by Lemma A.2 we have that $\frac{1}{JT}\sum_{it} |\tilde{Z}_{it}\mu_i' F_t| \xrightarrow{p} 0$ as $JT \to \infty$. The same argument as for $\tilde{\theta}_{YR}^{TSLS}$ for the first stage applies by noting that the same $\tilde{Z}_i t Z_{it}$ term appears. $\square$

1.5. Debiasing instrument Z

As noted in the main text debiasing the instrument is not a necessary condition for the consistent estimation of θ. In this section we note that while this is the case, debiasing the instrument can lead to better finite sample performance. In Lemma A.3 we derive a probability bound for the unobserved factor term in decomposition (6) for the cases in which the instrument Z or $\tilde{Z}$ is used. The Lemma shows that both lead to the same rate, but with a different constant. If Z is used the rate is multiplied by $\frac{1}{JT}Z'Z \xrightarrow{p} Q_Z$, the variation in the instrument Z. While, if $\tilde{Z}$ is used the rate is multiplied by Q, as described in Assumption 3, the variation in the instrument once the unobserved confounder U is projected out. Given that $Q \leq Q_Z$ by defition we expect debiasing the instrument to lead to a better finite sample rate. On the other hand, one might think that debiasing the instrument leads to a worse first stage. While this may be true, it is not guaranteed that using $\tilde{Z}$ will make the first stage worse. Under additional assumptions Lemma A.4 gives conditions under which using the instrument $\tilde{Z}$ leads to a weakly larger, but similar, first stage when we condition on the weights w. The intuition for the result is that $\tilde{Z}'Z$ appears in the first stage when the instrument Z is used versus $\tilde{Z}'\tilde{Z}$ in the case in which $\tilde{Z}$ is used. Suppose informally that $\tilde{Z} \simeq (I - P_U)Z$, where P_U is defined as in Assumption 3. Then, $\tilde{Z}'Z \simeq Z'(I - P_U)Z$ and $\tilde{Z}'\tilde{Z} \simeq Z'(I - P_U)'(I - P_U)Z = Z'(I - P_U)Z$ because $I - P_U$ is idempotent. Therefore, in both cases we do not expect the first stage to be very different.

In a simulation exercise, we plot the distribution of the empirical correlation between $\tilde{U}$, $\tilde{R}$ and Z and $\tilde{Z}$ in the context of our simulation design in section 6. As can be seen in Figure A.2 in the appendix simulation section, debiasing the instrument may matter a lot to reduce the finite sample correlation between the unobserved confounder and the instrument, but does not change the first stage significantly.

Lemma A.3 (Probability bounds on $Z'\tilde{U}$, $\tilde{Z}'U$ and $\tilde{Z}'\tilde{U}$). *Suppose in addition to Assumptions* $1-4$ *that* $\frac{1}{JT}Z'Z \xrightarrow{p} Q_Z$. *Then, under the assumptions of Theorem 4 it follows that as* $JT \to \infty$

$$\frac{1}{JT}\sum_{it}|Z_{it}\tilde{\mu}_i'F_t| \lesssim_P \sqrt{Q_Z} \times R(T,T_0,J),$$
$$\frac{1}{JT}\sum_{it}|\tilde{Z}_{it}\mu_i'F_t| \lesssim_P \sqrt{Q_Z} \times R(T,T_0,J),$$
$$\frac{1}{JT}\sum_{it}|\tilde{Z}_{it}\tilde{\mu}_i'F_t| \lesssim_P \sqrt{Q} \times R(T,T_0,J),$$

where $\lesssim_P$ *denotes bounded in probability and*

$$R(T,T_0,J) = \left(\frac{\bar{F}^2kc}{\xi}\right)\left(2c\frac{J}{\sqrt{T_0}}\sigma_\epsilon + (\bar{r}_1\bar{\sigma}_1 + \theta\bar{r}_2\bar{\sigma}_2)\left[\frac{1}{\sqrt{JT_0}} + JC^2\sqrt{\frac{T}{T_0}}\right]\right),$$

for $c = 1 + C$.

Proof. We start by noting that under the Assumptions of Theorem 2 we have that for the $JT \times 1$ vector $\tilde{U} = [\tilde{\mu}_i'F_t]_{it}$

$$\frac{1}{\sqrt{JT}}\mathbb{E}\|\tilde{U}\|_1 = \frac{1}{\sqrt{JT}}\sum_{i,t>T_0}\mathbb{E}|\tilde{\mu}_i'F_t| \le R(T,T_0,J),$$

where the inequality follows from the proof of Theorem 4 as the proof uses a bounding argument to account for $\tilde{Z}$. Next, define the $JT \times 1$ vector of instruments $Z = [Z_{it}]_{it}$ and observe that by the Cauchy-Schwartz inequality and the l_1-l_2 norm inequality we have that

$$\frac{1}{JT}|Z'\tilde{U}| \le \frac{1}{JT}\|Z\|_2\|\tilde{U}\|_2 \le \frac{1}{\sqrt{JT}}\|Z\|_2\frac{1}{\sqrt{JT}}\|\tilde{U}\|_1 \lesssim_P \frac{1}{\sqrt{JT}}\|Z\|_2 R(T,T_0,J),$$
$$\frac{1}{JT}|\tilde{Z}'\tilde{U}| \le \frac{1}{\sqrt{JT}}\|\tilde{Z}\|_2\frac{1}{\sqrt{JT}}\|\tilde{U}\|_2 \le \frac{1}{\sqrt{JT}}\|\tilde{Z}\|_2\frac{1}{\sqrt{JT}}\|\tilde{U}\|_1 \lesssim_P \frac{1}{\sqrt{JT}}\|\tilde{Z}\|_2 R(T,T_0,J),$$

where the last equality follows by applying Markov's inequality given that the expectation of the absolute value is bounded by rate R. The first result then follows by applying the continuous mapping theorem for the square root as $\frac{1}{JT}\|Z\|_2^2 \xrightarrow{p} Q_Z$. The second result follows by a similar argument to the proof of Theorem 2. As in the proof of Theorem 2, define the $T \times J$ matrix $\bar{Z}$, the $J \times J$ projection matrix U and the $J \times J$ weight matrix W. Then, we consider the following decomposition

$$\|\tilde{Z}\|_2 = \|(I-W)\bar{Z}\|_F = \|(I-P_U)Z + (P_U - \hat{W})Z\|_F^2.$$

In the proof of Theorem 2 we showed that $\frac{1}{JT}\|\tilde{Z}\|_F^2 \xrightarrow{p} Q$ as $JT \to \infty$. Therefore, by the continuous mapping theorem applied to the square root it follows that

$$\frac{1}{\sqrt{JT}}\|\tilde{Z}\|_2 \xrightarrow{p} \sqrt{Q},$$

which combined with the bound shows the first part of the proof. For the case of $\tilde{Z}'U$, we show that it is equivalent to the $Z'\tilde{U}$ case. We have that

$$\sum_{it} |\tilde{Z}_{it}\mu_i' F_t| = \sum_{it} |[Z'(I-\hat{W})'U]_{it}| = \|Z'(I-\hat{W})'U\|_1,$$

where $\|\cdot\|_1$ denotes the sum of the absolute value of the entries norm. Given that P_U is symmetric, a corollary of Lemma A.2 is that $\frac{1}{JT}\|(I-\hat{W})'U\|_1 \to 0$ in probability as $JT \to \infty$ since $\|\hat{W}' - P_U\|_F = \|(\hat{W} - P_U)'\|_F = \|\hat{W} - P_U\|_F$. Therefore, by Lemma A.2 we have that $\frac{1}{JT}\sum_{it}|\tilde{Z}_{it}\mu_i' F_t| \xrightarrow{p} 0$ as $JT \to \infty$ with the same rate $R(T, T_0, J)$. □

Lemma A.4 (First stage debiasing)**.** *Suppose that $Z_{it} = S_{it} + \mu_i' F_t$ and $A_{it} = 0$, where S_{it} is an i.i.d random variable such that $S_{it} \perp U_{it}$ and $\mathbb{E}[S_{it}] = 0$ and $\mathbb{E}[S_{it}^2] = \sigma_S^2$. Then, under the conditions of Theorem 2, for $\mathcal{W} = \Delta_J$, conditional on the weights w, it follows that as $JT \to \infty$*

$$\frac{1}{JT}\sum_{it} \tilde{R}_{it}\tilde{Z}_{it} \xrightarrow{p} \gamma\tilde{\xi},$$
$$\frac{1}{JT}\sum_{it} \tilde{R}_{it}Z_{it} \xrightarrow{p} \gamma\sigma_S^2,$$

where $\sigma_S^2 < \tilde{\xi} \leq 2\sigma_S^2$.

Proof. Under the assumptions, the first stage in the case in which the instrument is debiased and in the case in which it is not is given by, respectively

$$\frac{1}{JT}\sum_{it} \tilde{R}_{it}\tilde{Z}_{it} = \gamma\frac{1}{JT}\sum_{it} \tilde{S}_{it}^2 + o_p(1),$$
$$\frac{1}{JT}\sum_{it} \tilde{R}_{it}Z_{it} = \gamma\frac{1}{JT}\sum_{it} \tilde{S}_{it}S_{it} + o_p(1),$$

where the $o_p(1)$ terms are the terms involving $\tilde{Z}_{it}\tilde{\eta}_{it}$ and $Z_{it}\tilde{\eta}_{it}$ which converge to zero in probability by Lemma A.1, and the terms involving $\tilde{\mu}_i' F_t$ which converge to zero in probability by Theorem 1. Furthermore, by taking expectations with respect to S_{it} conditional on the weights $_i$ we have that

$$\mathbb{E}\left[\frac{1}{JT}\sum_{it}\tilde{S}_{it}S_{it}\right] = \frac{1}{JT}\sum_{it}\mathbb{E}[S_{it}^2] = \sigma_S^2,$$

$$\mathbb{E}\left[\frac{1}{JT}\sum_{it}\tilde{S}_{it}^2\right] = \frac{1}{JT}\sum_{it}\mathbb{E}[S_{it}^2](1+\|w_i\|_2^2) = \sigma_A^2\frac{1}{JT}\sum_{it}(1+\|w_i\|_2^2).$$

Given that the weights are in $\mathcal{W} = \Delta_J$, the simplex, it follows that $1/(J-1) \leq \|w_i\|_2^2 \leq 1$, so for all J and T we have that $\sigma_S^2 < \sigma_S^2\frac{1}{JT}\sum_{it}(1+\|w_i\|_2^2) \leq 2\sigma_S^2$. The result follows under an appropriate LLN. □

1.6. Proof of Theorem 4

In this section we provide a proof of Theorem 4 and an additional discussion in which we show that we can relax the conditioning on the weights by using a martingale representation of matching estimators as in Abadie and Imbens (2012).

Proof. We are interested in the following quantity for a given set of weights w and instrument Z:

$$\begin{aligned}\frac{\sqrt{JT}(\tilde{\theta}^{TSLS}-\theta)}{v_{JT}^w} &= \left(\frac{1}{JT}\sum_{it}\tilde{Z}_{it}\tilde{R}_{it}\right)^{-1}\frac{1}{v_{JT}^w\sqrt{JT}}\sum_{it}\tilde{Z}_{it}\left(\mu_i - \sum_{j\neq i}\hat{w}_{ij}^{SC}\mu_j\right)' F_t \\ &+ \left(\frac{1}{JT}\sum_{it}\tilde{Z}_{it}\tilde{R}_{it}\right)^{-1}\frac{1}{v_{JT}^w\sqrt{JT}}\sum_{it}\tilde{Z}_{it}\left(\epsilon_{it} - \sum_{j\neq i}\hat{w}_{ij}^{SC}\epsilon_{jt}\right),\end{aligned} \tag{A.6}$$

where the conditional variance is given by $v_{JT}^w = \frac{1}{\sqrt{JT}}\sum_{it} var(\tilde{Z}_{it}\tilde{\epsilon}_{it} \mid Z, w)$. First, we show that the bias term depending on $\mu_i' F_t$ in (A.6) is $o_p(1)$. The argument is the same as in the consistency theorem, but with a different rate. Note that from the proof of the consistency theorem (Theorem 3) that under Assumptions 1-4 we have that

$$\begin{aligned}\sum_{it}|F_t'(F_{T_0}F_{T_0}')^{-1}F_{T_0}\tilde{Y}_i^{T_0}| &\leq \left(\frac{\bar{F}^2 k}{\underline{\eta}T_0}\right)\sum_{it}|\sum_{t<T_0}\tilde{Y}_{it}| \\ &= \frac{T}{T_0}\left(\frac{\bar{F}^2 k}{\underline{\eta}}\right)\sum_{i,t<T_0}|\tilde{Y}_{it}|\end{aligned}$$

So, dividing by $\sqrt{JT}$ and using the bound on the pre-treatment mean absolute deviation as in

Theorem 2,

$$\begin{aligned}
\frac{1}{\sqrt{JT}}\sum_{it}|F_t'(F_{T_0}F_{T_0}')^{-1}F_{T_0}\tilde{Y}_i^{T_0}| &\leq \frac{\sqrt{T}}{T_0\sqrt{J}}\left(\frac{\bar{F}^2 k}{\underline{\eta}}\right)\sum_{i,t<T_0}|\tilde{Y}_{it}| \\
&\leq \frac{\sqrt{T}}{T_0\sqrt{J}}\left(\frac{\bar{F}^2 k}{\underline{\eta}}\right)\sqrt{JT_0}\bar{r}\bar{\sigma}_1(1+J) \\
&\lesssim \sqrt{\frac{T}{T_0}}(1+J)\bar{r}\bar{\sigma}_1.
\end{aligned}$$

Therefore, the first term is $o_p(1)$ when $\sqrt{\frac{T}{T_0}}(1+J)\bar{r}\bar{\sigma}_1 \to 0$. Similar argument follows for the R_{it} term.

Next, consider the sum in the second term in ($A.6$) and re-write it as follows

$$\begin{aligned}
\sum_{it}\tilde{Z}_{it}\tilde{\epsilon}_{it} &= \sum_{it}\tilde{Z}_{it}(\epsilon_{it} - \sum_{j\neq i}w_{ij}\epsilon_{jt}) \\
&= \sum_{it}\epsilon_{it}(\tilde{Z}_{it} - \sum_{j\neq i}\tilde{Z}_{jt}w_{ji}) \\
&= \sum_{it}\epsilon_{it}\tilde{\alpha}_{it},
\end{aligned}$$

where $\tilde{\alpha}_{it} = \tilde{Z}_{it} - \sum_{j\neq i}\tilde{Z}_{jt}w_{ji}$. Let $X_{it} \equiv \epsilon_{it}\tilde{\alpha}_{it}$. Given Assumption 4, it follows that conditional on Z and w, X_{it} is an i.i.d random variable with mean zero and variance $\sigma^2\tilde{\alpha}_{it}^2$. Indeed, we have that $\mathbb{E}[X_{it}|Z,w] = \tilde{\alpha}_{it}\mathbb{E}[\epsilon_{it}|Z,w] = 0$, and $\mathbb{V}(X_{it}|Z,w) = \mathbb{E}[\epsilon_{it}\tilde{\alpha}_{it}|Z,w] = \tilde{\alpha}_{it}^2\mathbb{E}[\epsilon_{it}|Z,w] = \tilde{\alpha}_{it}^2\sigma^2$. Next, let $s_{JT}^2 = \sum_{it} var(X_{it}|Z,w) = \sigma_\epsilon^2\|\tilde{\alpha}\|_2^2$ and consider the Lindeberg CLT condition for $\delta > 0$ conditional on Z and weights w

$$\begin{aligned}
\frac{1}{s_{JT}^2}\sum_{it}\mathbb{E}[X_{it}^2\mathbf{1}\{|X_{it}| > \delta s_{JT}\}] &= \frac{1}{\sigma_\epsilon^2\|\tilde{\alpha}\|_2^2}\sum_{it}\mathbb{E}[\epsilon_{it}^2\tilde{\alpha}_{it}^2\mathbf{1}\{|\epsilon_{it}| > \delta\frac{s_{JT}}{|\tilde{\alpha}_{it}|}\}] \\
&\leq \frac{1}{\sigma_\epsilon^2\|\tilde{\alpha}\|_2^2}\sum_{it}\mathbb{E}[\epsilon_{it}^2\tilde{\alpha}_{it}^2\mathbf{1}\{|\epsilon_{it}| > \delta\frac{s_{JT}}{max_{it}|\tilde{\alpha}_{it}|}\}] \\
&= \frac{1}{\sigma_\epsilon^2\|\tilde{\alpha}\|_2^2}\sum_{it}\tilde{\alpha}_{it}^2\mathbb{E}[\epsilon_{it}^2\mathbf{1}\{|\epsilon_{it}| > \delta\frac{s_{JT}}{max_{it}|\tilde{\alpha}_{it}|}\}] \\
&= \frac{1}{\sigma_\epsilon^2}\mathbb{E}[\epsilon_{it}^2\mathbf{1}\{|\epsilon_{it}| > \delta\frac{s_{JT}}{max_{it}|\tilde{\alpha}_{it}|}\}].
\end{aligned}$$

We start by showing that $\frac{max_{it}|\tilde{\alpha}_{it}|}{s_{JT}} \to 0$ as $JT \to \infty$ with high probability. Note that

$$max_{it}|\tilde{\alpha}_{it}| \leq c_z(1+C)(1+\|w^i\|_1),$$

where w^i denotes the vector of weights assigned to unit i across the $J-1$ synthetic controls for the other units. A simple bound on $\|w^i\|_1$ is given by $\|w^i\|_1 \leq CJ$, so $max_{it}|\tilde{\alpha}_{it}| \leq c_z(1+C)(1+CJ)$. For the denominator, we show that under our Assumptions as $JT \to \infty$

$$\frac{1}{JT}\|\tilde{\alpha}\|_2^2 \overset{p}{\not\to} 0.$$

We do so by re-writting the object of interest in matrix form for $J \times T$ matrices Z, $\tilde{Z}$ and $J \times J$ matrix of weights W. By the triangle inequality and the generalized Holder inequality for matrices, we have that

$$\begin{aligned}
\|\tilde{\alpha}\|_2^2 &= \|\tilde{Z} - W'\tilde{Z}\|_F^2 \\
&= \|(I-W)Z - W'(I-W)Z\|_F^2 \\
&\geq \|(I-W)Z\|_F^2 - \|W'(I-W)Z\|_F^2 \\
&\geq \|(I-W)Z\|_F^2 - \|W'(I-W)\|_\infty^2 \|Z\|_1^2.
\end{aligned}$$

By the proof of Theorem 2 we have that $\frac{1}{JT}\|(I-W)Z\|_F^2 \xrightarrow{p} Q > 0$ as $JT \to \infty$. For the second term in the last inequality, note that the instruments are bounded, so we have that $\|Z\|_1 \leq Jc_z$. Let P_U denote the $J \times J$ projection matrix for unobserved confounder $J \times T$ matrix U. It follows that, $P_U(I-P_U) = 0$. Then, consider

$$\|W'(I-W)\|_\infty^2 = \|W(I-W')\|_1^2 = \|W(I-W') - P_U(I-P_U)\|_1^2.$$

By Lemma A.2 we have that $\frac{1}{\sqrt{JT}}\|W - P_U\|_1 \xrightarrow{p} 0$ as $JT \to \infty$. It follows that $\frac{1}{JT}\|W(I-W')\|_1^2 = \|W(I-W') - P_U(I-P_U)\|_1^2 \xrightarrow{p} 0$. Therefore, as $J/T \to 0$ we have that $\frac{1}{JT}\|\tilde{\alpha}\|_2^2 \geq Q$ with high probability. It follows by an application of the continuous mapping theorem that with high probability,

$$\frac{max_{it}|\tilde{\alpha}_{it}|}{s_{JT}} \leq \frac{(1/\sqrt{JT})c_z(1+C)(1+CJ)}{1/\sqrt{JT}\sigma_\epsilon\|\tilde{\alpha}\|_2} \lesssim_P \sqrt{\frac{J}{T}}\frac{c_z C^2}{\sigma_\epsilon\sqrt{Q}}.$$

Therefore, as $\sqrt{J/T} \to 0$ we have that $\frac{max_{it}|\tilde{\alpha}_{it}|}{s_{JT}} \to 0$. In the case in which we condition on a weight sequence such that $\|w^i\|_1 \leq c_w$ for all i, it follows that $max_{it}|\tilde{\alpha}_{it}| \leq c_z(1+C)(1+c_w)$, so the requirement that $\sqrt{J/T} \to 0$ is no longer necessary and the rate holds as long as $JT \to \infty$.

The Lindeberg condition holds by the Dominated Convergence Theorem given that ϵ_{it} is $i.i.d$ and has bounded fourth moments. Therefore, we have that as $JT \to \infty$ and either $J/T \to \infty$ or for all i, $\|w^i\|_1 \leq c_w < \infty$,

$$\frac{1}{\sqrt{JT}v_{JT}^w}\sum_{it}\tilde{Z}_{it}\tilde{\epsilon}_{it} = \frac{1}{s_{JT}}\sum_{it}\tilde{Z}_{it}\tilde{\epsilon}_{it} \xrightarrow{d} N(0,1).$$

Combining this result with the first stage terms in (A.6) we have that under the same rates of convergence

$$\frac{\sqrt{JT}(\tilde{\theta}^{TSLS} - \theta)}{v^w_{JT}} \xrightarrow{d} Q^{-1}N(0,1),$$

where we used Slutsky's theorem and the fact that as $JT \to \infty$ $\frac{1}{JT}\sum_{it}\tilde{Z}_{it}\tilde{R}_{it} \xrightarrow{p} Q$. Define v^*_{JT} as the probability limit of the conditional variance estimate $v^w_{JT} = \frac{1}{\sqrt{JT}}\sum_{it} var(\tilde{Z}_{it}\tilde{\epsilon}_{it} \mid Z, w)$ as $JT \to \infty$. It follows by Slutsky's theorem that

$$\frac{\sqrt{JT}(\tilde{\theta}^{TSLS} - \theta)}{v_{JT}} \xrightarrow{d} N(0,1),$$

where $v_{JT} = Q^{-2}plim\frac{1}{\sqrt{JT}}\sum_{it} var(\tilde{Z}_{it}\tilde{\epsilon}_{it} \mid Z, w)$.

□

Discussion of weight conditioning Note that conditioning on the weights and instrument is equivalent to conditining on the pre-treatment period outcomes and treatments. Hence, it is possible consider a martingale representation as Abadie and Imbens (2012) do for matching estimators in which the information set is the pre-treatment period outcomes. Suppose for simplicity that we are in a setting where $R_{it} = 0$ for $t < T_0$ and $\mathcal{W} = \Delta_{J-1}$. Then, define the partial sums for a given time t as

$$S_{tJk} = \sum_{l=1}^{k} \tilde{Z}_{lt}\tilde{\epsilon}_{lt},$$

under our assumptions it follows that

$$\mathbb{E}[S_{tJk+1} \mid S_{tJ1}, \ldots, S_{tJk}] = \mathbb{E}[\tilde{Z}_{lt}\tilde{\epsilon}_{lt} \mid S_{tJ1}, \ldots, S_{tJk}] + S_{tJk} = S_{tJk},$$

where the condition expectation is zero given that conditional on the weights w under our error independence and partial instrument validity assumptions the instrument and the error term are uncorrelated when $t > T_0$ as shown in Lemma 1. Furthermore, define the martingale difference as

$$X_{tJk} = S_{tJk} - S_{tJk-1} = \tilde{Z}_{kt}\tilde{\epsilon}_{kt},$$

and the information set is given by the generated σ-algebra $\mathcal{F}_{tJk} = \sigma(\{Y_1^{T_0}, \ldots, Y_{k-1}^{T_0}, Z_{1t}, \ldots, Z_{k-1t}\})$ as the weights depend only on the outcome values in the pre-treatment period. Therefore, conditing on $\mathcal{F}_{tJk}$ is equivalent to conditiong on the weight vectors.

We can now apply the martingale CLT (Theorem 3.2, p. 59 from Hall and Heyde (1980)):

Theorem A.1 (Martingale CLT). *Let $\{S_{ni}, \mathcal{F}_{ni}, 1 \leq i \leq k_n, n \geq 1\}$ be a zero-mean, square-integrable* *martingale array with differences X_{ni} and let η^2 be an a.s. finite random variable. Suppose (1) a Lindeberg condition, for all $\varepsilon > 0$:*

$$\sum_i E\left(X_{ni}^2 \mathbf{1}\left\{|X_{ni}| > \varepsilon\right\} |\mathcal{F}_{n,i-1}\right) \xrightarrow{p} 0,$$

(2):

$$V_{nk_n}^2 = \sum_i E\left(X_{ni}^2 |\mathcal{F}_{n,i-1}\right) \xrightarrow{p} \eta^2,$$

and (3) the σ-fields are nested $\mathcal{F}_{n,i} \subset \mathcal{F}_{n+1,i}$ for $1 \leq i \leq k_n, n \geq 1$. Then:

$$S_{nk_n} = \sum_i X_{ni} \xrightarrow{d} Z,$$

where the random variable Z has characteristic function $E \exp\left(-\frac{1}{2}\eta^2 t^2\right)$.

Condition (3) is easy to check given our definition of $\mathcal{F}_{tJk}$. We start with condition (1) and consider $\frac{1}{\sqrt{JT}} X_{tJk}$. We will show point-wise convergence. Note that conditional on $\mathcal{F}_{tJk}$ by applying Holders inequality,

$$\mathbb{E}\left[\frac{X_{tJk}^2}{JT} \mathbf{1}\left\{|X_{tJk}| > \varepsilon\right\}\right] \leq \frac{1}{JT} \mathbb{E}\left[X_{tJk}^4\right]^{1/2} P\left(\left\{|X_{tJk}| > \sqrt{JT}\varepsilon\right\}\right)^{1/2}.$$

The second term can be further bounded by applying Chebyshev's inequality and under Assumptions 1-3

$$\begin{aligned} P\left(\left\{|X_{tJk}| > \sqrt{JT}\varepsilon\right\}\right) &\leq \frac{var(\tilde{\epsilon}_{kt}|w) c_z^2}{JT\varepsilon^2} \\ &\leq \frac{\sigma^2 c_z^2}{T\varepsilon^2}, \end{aligned}$$

where the conditional variance is bounded above by the sum of the J variances. The first expectation term can be bounded by noting that the instruments are bounded and under the assumption of bounded fourth moments of the error term

$$\begin{aligned} \mathbb{E}\left[X_{tJk}^4\right] &\leq c_z^4 \mathbb{E}\left[\tilde{\epsilon}_{kt}^4\right] \\ &\leq c_z^4 \mathbb{E}\left[\sum_i \epsilon_{it}^4\right] \\ &= c_z^4 J \mathbb{E}\left[\epsilon_{it}^4\right]. \end{aligned}$$

Combining the two bounds we get that for $X_{tJk}/\sqrt{JT}$,

$$\sum_{kt} E\left((X_{tJk}/\sqrt{JT})^2 \mathbf{1}\left\{|X_{tJk}/\sqrt{JT}| > \varepsilon\right\}|\mathcal{F}_{tJk-1}\right) \leq JT\frac{\sigma\sqrt{m_4}c_z^3\sqrt{J}}{JT\sqrt{T}\varepsilon} \lesssim \sqrt{\frac{J}{T}},$$

where $\mathbb{E}\left[\epsilon_{it}^4\right] \leq m_4$. Hence, as $\sqrt{\frac{J}{T}} \to 0$ Lindeberg's condition (1) is satisfied point-wise. Next, note that the variance term in condition (2) is bounded in probability and not $o_p(1)$ by the proof of Theorem 4. Hence, all conditions for the martingale Lindeberg CLT are satisfied and the normality results extends to the case in which we do not directly condition on the weights.

1.7. Results for event study designs

To derive theoretical results for the event study designs using the SIV debiased values, we impose additional assumptions on the instrument and the reduced form equation.

Assumption 5 (Event study design)**.** *The outcome of interest follows*

$$Y_{it} = \theta_t Z_i + \mu_i' F_t + \epsilon_{it},$$

where θ_t is the time varying parameter of interest satisfying $\theta_t = 0$ for $t \leq T_0$ and the instrument satisfies $Z_i \perp \epsilon_{it}, \eta_{it}$, $|Z_i| \leq c_z$, for all i, t, and, as $J \to \infty$,

$$\frac{1}{J}Z'(I - P_U)Z \xrightarrow{p} Q_e > 0,$$

where Z is the $J \times 1$ vector containing Z_i for all i and $P_U = U(U'U)^{-1}U'$ is the projection matrix for the $J \times k$ matrix of factor loadings U.

Note, that while Assumption 5 abstracts away from the instrument equation of the triangular design in Assumption 1, it can directly be reconciled with our design. Consider plugging in the instrument equation from Assumption 1 into the outcome equation and setting $A_{it} = 0$ such that

$$Y_{it} = \theta(\gamma Z_{it} + \eta_{it}) + \mu_i' F_t + \epsilon_{it}.$$

Furthermore, suppose we have a factor instrument, as is the case in the event study designs, and $Z_{it} = Z_i \times H_t$ with $H_t = 0$ for $t \leq T_0$, then we have that $\theta_t \equiv \gamma\theta H_t$ and $\theta_t = 0$ for $t \leq T_0$. Furthermore, the error term satisfies that $\epsilon_{it} + \theta\eta_{it} \perp Z_i$ as required by Assumption 5.

Under Assumption 5 we will show that the least-squares estimator for θ_t based on equation

$$\tilde{Y}_{it} = \sum_{k \neq T_0} \theta_k (\mathbb{1}\{t = k\} \times \tilde{Z}_i) + \epsilon_{it}, \tag{A.7}$$

recovers the dynamic effect θ_t under the same additional assumptions as Theorem 2. The SIV least squares event study estimator solves the following program for $\theta \in [\theta_1, \ldots, \theta_{T_0-1}, \theta_{T_0+1}, \ldots, \theta_T]$,

$$\min_\theta \sum_{it} (\tilde{Y}_{it} - \sum_{k \neq T_0} \theta_k (\mathbb{1}\{t = k\} \times \tilde{Z}_i))^2.$$

Taking FOC with respect to θ_l for some $l \neq T_0$ we have that

$$-2 \sum_{it} (\tilde{Y}_{it} - \sum_{k \neq T_0} \theta_k (\mathbb{1}\{t = k\} \times \tilde{Z}_i)) \mathbb{1}\{t = l\} \tilde{Z}_i = 0$$

$$\sum_i (\tilde{Y}_{il} - \theta_l \tilde{Z}_i) \tilde{Z}_i = 0,$$

as $\mathbb{1}\{t = k\}\mathbb{1}\{t = l\} = 1$ if and only if $k = l$. Hence, the SIV event study estimator is given by

$$\tilde{\theta}_l = \left(\sum_i \tilde{Z}_i^2 \right)^{-1} \sum_i \tilde{Z}_i \tilde{Y}_{il}.$$

Similarly to the SIV estimator for θ, we can decompose the event study estimator as follows

$$\tilde{\theta}_l = \theta_l + \left(\sum_i \tilde{Z}_i^2 \right)^{-1} \sum_i \tilde{Z}_i \tilde{\mu}_i' F_l + \left(\sum_i \tilde{Z}_i^2 \right)^{-1} \sum_i \tilde{Z}_i \tilde{\epsilon}_{il}. \tag{A.8}$$

In what follows, we prove a consistency result (Theorem A.2) and asymptotic normality result (Theorem A.3) for the event study estimator under the similar conditions as the results in the main text (Theorem 2 and Theorem 4). The main difference is that our results for the event study estimator require the number of units J and pre-treatment periods T_0 to grow to infinity, whether the results in the main text are valid for fixed J.

Theorem A.2 (Event study estimates consistency)**.** *Under Assumption 5 and the additional assumptions of Theorem 2, it follows that as $J \to \infty$ and $\bar{r}_1 \bar{\sigma}_1 \sqrt{\frac{J}{T_0}} \to 0$*

$$\tilde{\theta}_l \xrightarrow{p} \theta_l$$

for $l \in \{1, \ldots, T_0 - 1, T_0 + 1, \ldots, T\}$.

Proof. Start by considering the terms in A.8 that involve $\tilde{\epsilon}_{il}$. Under Assumption 3 and $Z_i \perp \epsilon_{it}$ it follows by Lemma A.1 for $t = l$ that as $J \to \infty$

$$\frac{1}{J}\sum_i \tilde{Z}_i \epsilon_{il} \xrightarrow{p} 0.$$

Next, we consider the term involving $\tilde{\mu}_i' F_l$. Observe that under Assumption 5 we have that for $t \leq T_0$

$$\tilde{Y}_{it} = \tilde{\mu}_i' F_t + \tilde{\epsilon}_{it},$$

so we proceed as in the proof of Theorem 1 by projecting out the common factors and applying the triangle inequality.

$$\begin{aligned}
\left|\sum_i \tilde{Z}_i \tilde{\mu}_i' F_l\right| &= \left|\sum_i \tilde{Z}_i F_l'(F_{T_0}F_{T_0}')^{-1}F_{T_0}(\tilde{Y}_i^{T_0} - \tilde{\epsilon}_{it})\right| \\
&\leq \sum_i |\tilde{Z}_i F_l'(F_{T_0}F_{T_0}')^{-1}F_{T_0}\tilde{\epsilon}_i| + \sum_i |\tilde{Z}_i F_l'(F_{T_0}F_{T_0}')^{-1}F_{T_0}\tilde{Y}_i^{T_0}| \\
&\leq (1+C)c_z\left(\sum_i |F_l'(F_{T_0}F_{T_0}')^{-1}F_{T_0}\tilde{\epsilon}_i| + \sum_i |F_l'(F_{T_0}F_{T_0}')^{-1}F_{T_0}\tilde{Y}_i^{T_0}|\right).
\end{aligned}$$

Applying the bounds from the proof of Theorem 1 we have that

$$\frac{1}{J}\sum_i \mathbb{E}[|F_l'(F_{T_0}F_{T_0}')^{-1}F_{T_0}(\epsilon_{iT_0} - \epsilon_{-iT_0}' w_i)|] \leq 2(1+C)\left(\frac{\bar{F}^2 k}{\underline{\xi}}\right)\sqrt{\frac{J}{T_0}}\sigma_\epsilon,$$

and for the term involving $\tilde{Y}$ we get that

$$\frac{1}{J}\sum_i \mathbb{E}|F_t'(F_{T_0}F_{T_0}')^{-1}F_{T_0}\tilde{Y}_i^{T_0}| \leq \left(\frac{\bar{F}^2 k}{\underline{\eta}}\right)\mathbb{E}\left[\frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{Y}_{it}|\right].$$

Given that the bounds depend on the same objects as in the proof of Theorem 2, under the stable rank Assumption 4, it follows that as $J \to \infty$ and $\bar{r}_1\bar{\sigma}_1\sqrt{\frac{J}{T_0}} \to 0$

$$\frac{1}{J}\sum_i \tilde{Z}_i \tilde{\mu}_i{}' F_l \xrightarrow{p} 0.$$

Applying Lemma A.2 for the $t = l$ case, it follows that $\frac{1}{J}\sum_i \tilde{Z}_i^2 \xrightarrow{p} Q_e > 0$ under the same regime as $J \to \infty$, which completes the consistency result. $\square$

Theorem A.3 (Asymptotic normality for event study estimates)**.** *Under Assumption 8, and the additional assumptions of Theorem 2, conditional on weights w and instruments Z_i, if* $\sqrt{\frac{1}{T_0}}(1 +$

$J)\bar{r}_1\bar{\sigma}_1 \to 0$ and $\frac{1}{\sqrt{J}}\max_i \sum_{j\neq i}|w_{ji}| \to 0$, then as $J\to\infty$

$$\frac{\sqrt{J}(\tilde{\theta}_l-\theta_l)}{v_J} \xrightarrow{d} (Q_e)^{-1}\times N(0,1).$$

where $v_J^2 = \frac{1}{J}\sum_i var(\tilde{Z}_i\tilde{\epsilon}_i \mid Z,w) = \frac{1}{J}\sum_i \sigma_\epsilon^2\tilde{\alpha}_i^2$ and $\tilde{\alpha}_i = \tilde{Z}_i - \sum_{j\neq i}\tilde{Z}_j w_{ji}$.

Proof. We start by decomposing the estimator of interest

$$\begin{aligned}\frac{\sqrt{J}(\tilde{\theta}_l-\theta_l)}{v_J^w} &= \left(\frac{1}{J}\sum_i \tilde{Z}_i^2\right)^{-1}\frac{1}{v_{Jl}^w\sqrt{J}}\sum_i \tilde{Z}_i^2\left(\mu_i-\sum_{j\neq i}\hat{w}_{ij}^{SC}\mu_j\right)' F_l \\ &\quad + \left(\frac{1}{J}\sum_i \tilde{Z}_i^2\right)^{-1}\frac{1}{v_J^w\sqrt{J}}\sum_i \tilde{Z}_i\left(\epsilon_i-\sum_{j\neq i}\hat{w}_{ij}^{SC}\epsilon_{jl}\right),\end{aligned} \tag{A.9}$$

where $(v_J^w)^2 = \frac{1}{J}\sum_i var(\tilde{Z}_i\tilde{\epsilon}_i \mid Z,w)$. The consistency of the factor term follows from a similar bounding argument as in the proof of Theorem 2. Given Assumption 3, we have that

$$\begin{aligned}\frac{1}{\sqrt{J}}\sum_i |F_t'(F_{T_0}F_{T_0}')^{-1}F_{T_0}\tilde{Y}_i^{T_0}| &\leq \frac{1}{T_0\sqrt{J}}\left(\frac{\bar{F}^2k}{\underline{\eta}}\right)\sum_{i,t<T_0}|\tilde{Y}_{it}| \\ &\leq \frac{1}{T_0\sqrt{J}}\left(\frac{\bar{F}^2k}{\underline{\eta}}\right)\sqrt{JT_0}\bar{r}\bar{\sigma}_1(1+J) \\ &\lesssim \sqrt{\frac{1}{T_0}}(1+J)\bar{r}\bar{\sigma}_1.\end{aligned}$$

Given that under Assumption 5 as $J\to\infty$ we have that $\frac{1}{J}\sum_i \tilde{Z}_i^2 \xrightarrow{p} Q_e > 0$, it follows that the first term in A.9 is $o_p(1)$ when $\sqrt{\frac{1}{T_0}}(1+J)\bar{r}_1\bar{\sigma}_1 \to 0$ as $T_0, J\to\infty$.

We proceed by showing that the second term A.9 converges in distribution to a normal random variable.

$$\begin{aligned}\sum_i \tilde{Z}_i\tilde{\epsilon}_{il} &= \sum_i \tilde{Z}_i(\epsilon_{il}-\sum_{j\neq i}w_{ij}\epsilon_{jl}) \\ &= \sum_i \epsilon_{il}(\tilde{Z}_i-\sum_{j\neq i}\tilde{Z}_j w_{ji}) \\ &= \sum_i \epsilon_{il}\tilde{\alpha}_i,\end{aligned}$$

where $\tilde{\alpha}_i = \tilde{Z}_i - \sum_{j\neq i}\tilde{Z}_j w_{ji}$. As in the proof of Theorem 4, we show that the Lindeberg CLT condition holds as $J\to\infty$. Let $X_{il}\equiv \epsilon_{il}\tilde{\alpha}_i$. Given Assumption 5, it follows that conditional on Z and w, X_{il} is an i.i.d random variable with mean zero and variance $\sigma^2\tilde{\alpha}_i^2$. Indeed, we have that

$\mathbb{E}[X_{il}|Z,w] = \tilde{\alpha}_i \mathbb{E}[\epsilon_{il}|Z,w] = 0$, and $\mathbb{V}(X_{il}|Z,w) = \mathbb{E}[\epsilon_{il}\tilde{\alpha}_i|Z,w] = \tilde{\alpha}_i^2 \mathbb{E}[\epsilon_{il}|Z,w] = \tilde{\alpha}_i^2\sigma^2$. Next, let $s_J^2 = \sum_{it} var(X_{il}|Z,w) = \sigma_\epsilon^2\|\tilde{\alpha}\|_2^2$ and consider the Lindeberg CLT condition for $\delta > 0$ conditional on Z and weights w.

$$\begin{aligned}\frac{1}{s_J^2}\sum_i \mathbb{E}[X_{il}^2\mathbf{1}\{|X_{il}| > \delta s_J\}] &= \frac{1}{\sigma_\epsilon^2\|\tilde{\alpha}\|_2^2}\sum_i \mathbb{E}[\epsilon_{il}^2\tilde{\alpha}_i^2\mathbf{1}\{|\epsilon_{il}| > \delta\frac{s_J}{|\tilde{\alpha}_i|}\}] \\ &= \frac{1}{\sigma_\epsilon^2}\mathbb{E}[\epsilon_{il}^2\mathbf{1}\{|\epsilon_{il}| > \delta\frac{s_J}{\max_i|\tilde{\alpha}_i|}\}].\end{aligned}$$

Under our assumptions we have that $\frac{1}{\sqrt{J}}\max_i|\tilde{\alpha}_i| \to 0$ as $J \to \infty$. Therefore, it suffices to show that $\frac{1}{\sqrt{J}}s_J$ converges to a constant to guarantee the Lindeberg condition, given that we assume bounded fourth moments. A similar argument to the proof of Theorem 4 shows that with high probability as $J \to \infty$

$$\frac{1}{J}\|\tilde{\alpha}\| \geq Q_e > 0,$$

so by an application of of the continuous mapping theorem we have as $J \to \infty$,

$$\frac{s_J}{\max_i|\tilde{\alpha}_i|} \to 0.$$

The asymptotic normality result then follows by an application of Slutsky's Theorem and the Lindeberg CLT. □

Event study variance estimation The conditional variance v_J^2 can be estimated directly from the data using the plug-in estimator for idiosyncratic shock variance $\hat{\sigma}_\epsilon^2$. In our applications, we let the variance of the error term vary with l, and we estimate the variance as follows

$$\tilde{\sigma}_{l,SIV}^2 = \frac{\hat{\sigma}_{\epsilon,l}^2\|\tilde{\alpha}\|_2^2}{\sum_i \tilde{Z}_i^2},$$

where $\hat{\sigma}_{\epsilon,l}^2 = \frac{1}{J-1}\sum_i \hat{\epsilon}_{il}^2$, where $\hat{\epsilon}_{il}^2$ is the residual for the event study regression for time period parameter θ_l. Standard errors and asymptotically valid confidence intervals can be constructed using this variance estimator under the same assumptions as Theorem A.3, if additionally $\mathbb{E}[\epsilon_{il}^4] < \infty$ for all l. In the event study figures, e.g. Figure 4, the 95% confidence intervals vary with time because we use the time-varying variance estimator.

1.8. Projected and ensemble estimators

The bound for the projected estimator can be derived in the same way as for the SIV estimator in the proof of Theorem 1. We give a sketch of the proof. The only difference for the projected estimator is that we now have that

$$\tilde{\mu}_i^P = (F_{T_0} F_{T_0}')^{-1} F_{T_0} (\tilde{Y}_i^{P,T_0} - \theta \tilde{R}_i^{P,T_0} - \tilde{\epsilon}_{it}^P).$$

Where the $\tilde{\epsilon}_{it}^P$ denotes the projection into the instrument space Z_i. Given that $Z_i \perp \epsilon_{it}$ and ϵ_{it} is mean zero, it will mean that with high probability the weights do not depend on the ϵ_{it} terms in the pre-treatment period. Hence, removing the contribution of the J other units in the rate.

Under Assumptions 1-4, the same bounds apply for $\mathbb{E}\left[\frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{Y}_{jt}^P|\right] + \theta\mathbb{E}\left[\frac{1}{JT_0}\sum_{i,t\leq T_0}|\tilde{R}_{jt}^P|\right]$ for the projected estimator than for the SIV up to an additional constant depending on Q (from the projection operator $Z'(Z'Z)^{-1}Z$). This highlights that in cases in which the instrument does not explain a lot of the variation in U then projecting into the instrument space makes matching U harder.

Therefore, the projected SIV is also a consistent estimator of θ under the same conditions as the SIV. Hence, for any $\alpha \in (0,1)$ the ensemble estimator is consistent.

Aggregation estimator As an alternative to the projected estimator when Z_{it} does not follow a factor structure, we consider an aggregation estimator that matches on the aggregated timeseries in the pre-period. The Aggregation estimator is computed equivalently to the projected estimator except that the synthetic control weigths are computed as follows

1. Let $Q_i = \sum_{t<T_0} Z_i Y_{it}$ for each i.
2. Match the aggregated values for each i

$$w_j^A \in \text{argmin}_{w\in\Delta^{J-1}} \|Q_i - Q_{-i}'w\|^2$$

3. Compute the aggregated TSLS estimator $\tilde{\theta}^A$ using the w^A weights.

In the appendix additional simulation section we provide results for the aggregation estimator and show that it performs badly in settings in which the time series component is important.

1.9. Randomization Inference

An alternative to the permutation based test described in section 5 is a test based on randomization inference in the spirit of Imbens and Rosenbaum (2005). Instead of considering the differences in effects across time periods we now fix outcomes Y and consider different assignment distributions for the instrument-treatment pairs (R, Z) across units. The intuition is that under a uniform assignment distribution of (R, Z) we should not see an effect on the outcome Y. We can construct a permutation t-statistic and p-value using the following procedure.

1. In the pre-period compute the SC weights and generate the debiased quantities.

2. Define the set of permutations of the J units: $\mathcal{P}(J)$.

3. For a given permutation $\pi \in \mathcal{P}(J)$, compute

$$\tilde{\theta}_\pi = \left(\sum_{it} \tilde{Z}_{\pi(i)t}\tilde{R}_{\pi(i)t}\right)^{-1} \sum_{it} \tilde{Z}_{\pi(i)t}\tilde{Y}_{it},$$

   where we permute the individuals for Z and R but not Y.

4. p-value:

$$\hat{p} = \frac{1}{\mathcal{P}(J)} \sum_{\pi \in \mathcal{P}(J)} P(\tilde{\theta}_\pi \geq \tilde{\theta}_{TSLS})),$$

   where absolute values should be used for a two-sided test.

This test will be a valid test for the null $H_0 : \theta = 0$ under the assumption of exchangeability across units of $\{\mu_i, \epsilon_{it}, \eta_{it}$. The reason why we do not implement this test for our main empirical example is that we do not believe that μ_i, which is correlated with the distance share Z_i is exchangeable across units. It should also be noted that in general $\mathcal{P}(J)$ might be very large and in practice the p-values should be computed by randomly sampling from $\mathcal{P}(J)$.

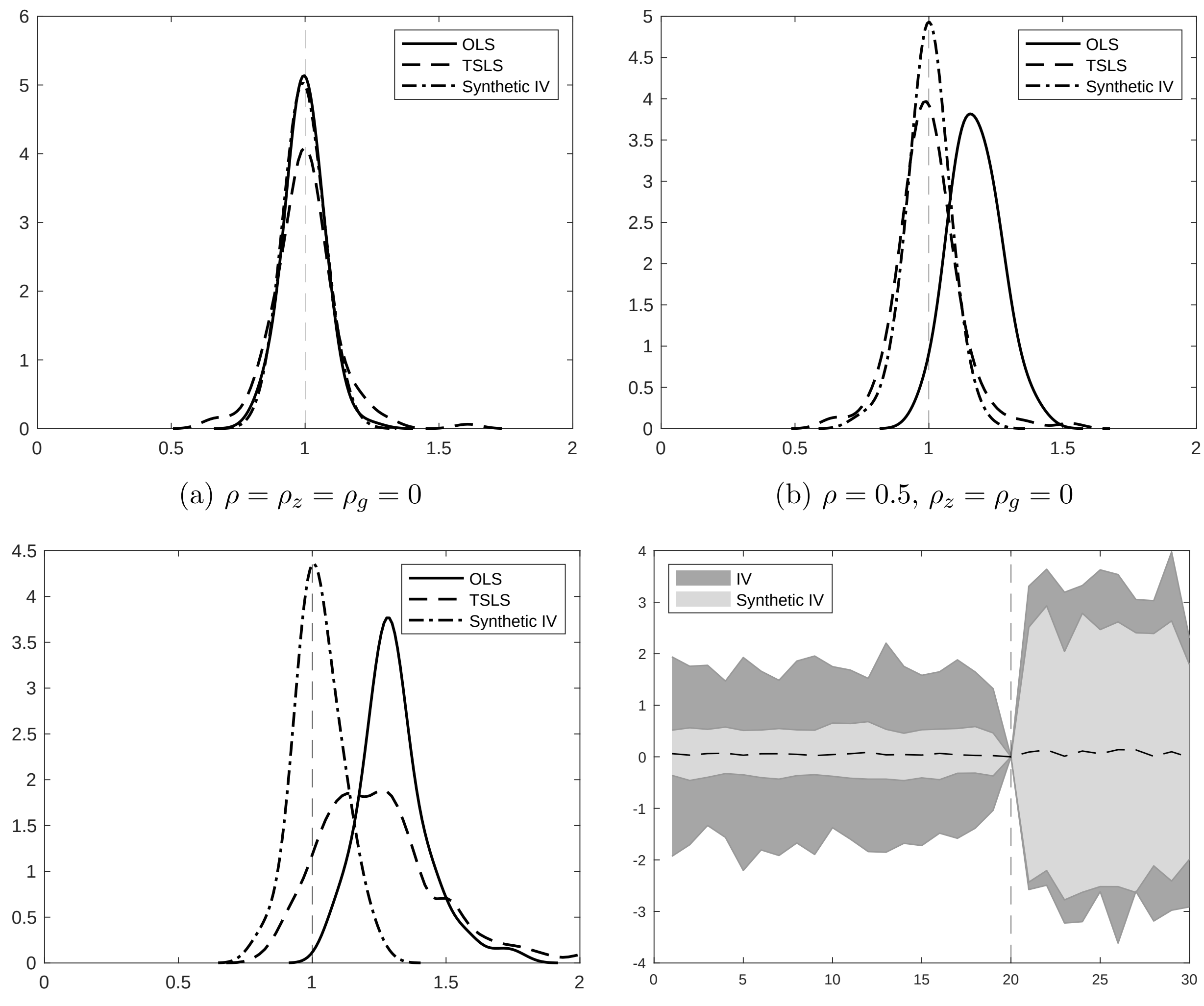
1.10. Additional simulations

In this section, we consider the same simulation design as in section 6 but under different parameters. In particular, we consider a setting with a weaker instrument and a smaller signal variance. Figure A.1.1 replicates the first five panels of Figure 5 for a simulation design with parameters $\beta = \gamma = 1$, $k = 1$, $T = 30$, $T_0 = 20$, $J = 20$, $\sigma_\epsilon = 0.5$, $\kappa = 0.5$, $\sigma_\eta = \sigma_z = \sigma_g = 1$.

Table A.1 replicates Table 1 for the simulation design of Figure A.1.1 and Table A.3 for the design with $T_0 = 10$. In all cases we find that the SIV estimator, projected SIV and ensemble estimator outperform the OLS and TSLS estimators. Furthermore, the performance of the SIV whe the number of pre-treatment periods is halved remains good. Table A.2 evaluates the coverage of the 95% confidence intervals for the synthetic IV using $\hat{v}_{JT_1}$ of the true parameter $\theta = 1$ for different correlation settings and post-treatment periods. We find that in settings in which the OLS and TSLS are unbiased the synthetic IV exhibits a slight over-coverage, in the well behaved settings with

Figure A.1.1: Model comparison in simulations

Note: Panels (a)-(c) display kernel density plots for TWFE OLS, TWFE TSLS and the synthetic IV. Panel (d) shows simulated event study estimates as in Figure 2 panel (d) with 95% confidence bands for $\rho = \rho_z = \rho_g = 0.5$*. Simulations are done over 1000 iterations with the following parameters:* $\beta = \gamma = 1$, $k = 1$, $T = 30$, $T_0 = 20$, $J = 20$, $\sigma_\epsilon = 0.5$, $\kappa = 0.5$, $\sigma_\eta = \sigma_z = \sigma_g = 1$.



(a) $\rho = \rho_z = \rho_g = 0$

(b) $\rho = 0.5$, $\rho_z = \rho_g = 0$

(c) $\rho = \rho_z = \rho_g = 0.5$

(d) Event study estimates

moderate noise and correlation the coverage is good, and in high correlation settings, as expected, we report under-coverage.

Table A.1: Simulations for different $r = \rho = \rho_z = \rho_g$ and σ_ϵ.

	r=0.5				r=0.7				r=0.9			
	Mean	Var	Bias	MSE	Mean	Var	Bias	MSE	Mean	Var	Bias	MSE
					$\sigma_\epsilon = 0.5$							
OLS (TWFE)	1.31	0.02	0.31	0.11	1.50	0.02	0.50	0.27	1.73	0.01	0.73	0.55
TSLS (TWFE)	1.26	0.07	0.26	0.13	1.51	0.08	0.51	0.34	1.83	0.06	0.83	0.74
SIV	1.02	0.01	0.02	0.01	1.05	0.02	0.05	0.02	1.19	0.04	0.19	0.07
projected SIV	0.92	0.03	-0.08	0.04	0.95	0.04	-0.05	0.05	1.11	0.07	0.11	0.08
Agg. SIV	1.23	0.08	0.23	0.13	1.46	0.08	0.46	0.29	1.80	0.04	0.80	0.68
SIV + projected	1.01	0.01	0.01	0.01	1.03	0.02	0.03	0.02	1.15	0.04	0.15	0.06
SIV + Agg.	1.03	0.01	0.03	0.01	1.07	0.02	0.07	0.02	1.21	0.04	0.21	0.08
SIZ Z	1.07	0.02	0.07	0.02	1.15	0.03	0.15	0.05	1.43	0.03	0.43	0.21
					$\sigma_\epsilon = 1$							
OLS (TWFE)	1.38	0.02	0.38	0.16	1.60	0.02	0.60	0.38	1.86	0.02	0.86	0.76
TSLS (TWFE)	1.26	0.07	0.26	0.14	1.50	0.08	0.50	0.34	1.82	0.06	0.82	0.74
SIV	1.03	0.01	0.03	0.01	1.07	0.03	0.07	0.03	1.26	0.05	0.26	0.12
projected SIV	0.90	0.05	-0.10	0.06	0.94	0.06	-0.06	0.07	1.14	0.08	0.14	0.10
Agg. SIV	1.22	0.07	0.22	0.12	1.47	0.08	0.47	0.30	1.80	0.04	0.80	0.69
SIV + projected	1.01	0.01	0.01	0.01	1.03	0.03	0.03	0.03	1.21	0.05	0.21	0.10
SIV + Agg.	1.04	0.01	0.04	0.02	1.10	0.03	0.10	0.04	1.30	0.05	0.30	0.14
SIZ Z	1.08	0.02	0.08	0.03	1.19	0.03	0.19	0.07	1.50	0.03	0.50	0.28
					$\sigma_\epsilon = 2$							
OLS (TWFE)	1.48	0.02	0.48	0.26	1.74	0.03	0.74	0.58	2.05	0.02	1.05	1.12
TSLS (TWFE)	1.26	0.08	0.26	0.14	1.50	0.09	0.50	0.34	1.82	0.07	0.82	0.74
SIV	1.05	0.03	0.05	0.03	1.12	0.04	0.12	0.06	1.37	0.07	0.37	0.21
projected SIV	0.87	0.08	-0.13	0.09	0.93	0.10	-0.07	0.10	1.21	0.10	0.21	0.15
Agg. SIV	1.22	0.08	0.22	0.12	1.46	0.09	0.46	0.30	1.80	0.05	0.80	0.68
SIV + projected	1.01	0.03	0.01	0.03	1.06	0.05	0.06	0.05	1.29	0.08	0.29	0.16
SIV + Agg.	1.07	0.03	0.07	0.03	1.16	0.05	0.16	0.07	1.43	0.07	0.43	0.26
SIZ Z	1.10	0.03	0.10	0.04	1.24	0.05	0.24	0.10	1.58	0.04	0.58	0.38
					$\sigma_\epsilon = 4$							
OLS (TWFE)	1.63	0.04	0.63	0.43	1.95	0.04	0.95	0.94	2.31	0.04	1.31	1.75
TSLS (TWFE)	1.25	0.09	0.25	0.15	1.49	0.11	0.49	0.35	1.81	0.08	0.81	0.74
SIV	1.08	0.05	0.08	0.05	1.19	0.07	0.19	0.11	1.49	0.10	0.49	0.34
projected SIV	0.85	0.13	-0.15	0.15	0.95	0.15	-0.05	0.15	1.30	0.15	0.30	0.24
Agg. SIV	1.20	0.10	0.20	0.14	1.45	0.11	0.45	0.31	1.79	0.07	0.79	0.69
SIV + projected	1.01	0.05	0.01	0.05	1.10	0.08	0.10	0.09	1.41	0.11	0.41	0.28
SIV + Agg.	1.11	0.05	0.11	0.06	1.24	0.07	0.24	0.13	1.56	0.09	0.56	0.40
SIZ Z	1.13	0.06	0.13	0.07	1.30	0.07	0.30	0.16	1.65	0.06	0.65	0.48
					$\sigma_\epsilon = 8$							
OLS (TWFE)	1.83	0.06	0.83	0.75	2.24	0.08	1.24	1.61	2.68	0.09	1.68	2.91
TSLS (TWFE)	1.24	0.11	0.24	0.17	1.49	0.13	0.49	0.37	1.80	0.11	0.80	0.76
SIV	1.12	0.09	0.12	0.11	1.27	0.12	0.27	0.19	1.61	0.13	0.61	0.50
projected SIV	0.88	0.22	-0.12	0.24	1.01	0.24	0.01	0.24	1.42	0.25	0.42	0.42
Agg. SIV	1.19	0.15	0.19	0.19	1.43	0.15	0.43	0.34	1.78	0.11	0.78	0.72
SIV + projected	1.05	0.10	0.05	0.10	1.18	0.13	0.18	0.17	1.53	0.16	0.53	0.44
SIV + Agg.	1.15	0.10	0.15	0.12	1.32	0.12	0.32	0.23	1.66	0.11	0.66	0.55
SIZ Z	1.16	0.09	0.16	0.12	1.36	0.11	0.36	0.24	1.71	0.09	0.71	0.59

Table A.2: $\mathrm{T}_0 = 20, J = 20, \sigma_\epsilon = 0.5, \sigma_z = 1, \sigma_{other} = 0.5, \kappa = 0.5.$

Coverage $\alpha = 0.05$			
	T=30	T=40	T=50
$\rho = \rho_g = \rho_z = 0.0$	0.981	0.962	0.952
$\rho = \rho_g = \rho_z = 0.3$	0.976	0.944	0.96
$\rho = \rho_g = \rho_z = 0.5$	0.960	0.945	0.923
$\rho = \rho_g = \rho_z = 0.7$	0.904	0.808	0.792

Table A.3: Simulations for $T_0 = 10$

	r=0.5				r=0.7				r=0.9			
	Mean	Var	Bias	MSE	Mean	Var	Bias	MSE	Mean	Var	Bias	MSE
					$\sigma = 0.5$							
OLS (TWFE)	1.29	0.02	0.29	0.10	1.48	0.02	0.48	0.25	1.71	0.01	0.71	0.52
TSLS (TWFE)	1.23	0.05	0.23	0.11	1.46	0.07	0.46	0.29	1.78	0.06	0.78	0.68
SIV	1.01	0.01	0.01	0.01	1.07	0.02	0.07	0.02	1.25	0.04	0.25	0.10
projected SIV	0.93	0.04	-0.07	0.04	0.99	0.05	-0.01	0.05	1.24	0.36	0.24	0.41
Agg. SIV	1.23	0.06	0.23	0.12	1.47	0.07	0.47	0.30	1.79	0.05	0.79	0.68
SIV + projected	0.94	0.04	-0.06	0.04	0.99	0.05	-0.01	0.05	1.24	0.35	0.24	0.41
SIV + Agg.	1.23	0.06	0.23	0.12	1.47	0.07	0.47	0.29	1.79	0.05	0.79	0.67
SIZ Z	1.05	0.02	0.05	0.02	1.16	0.03	0.16	0.05	1.49	0.03	0.49	0.27
					$\sigma = 1$							
OLS (TWFE)	1.36	0.02	0.36	0.15	1.59	0.02	0.59	0.36	1.84	0.01	0.84	0.73
TSLS (TWFE)	1.22	0.06	0.22	0.11	1.46	0.07	0.46	0.29	1.78	0.06	0.78	0.67
SIV	1.02	0.02	0.02	0.02	1.11	0.03	0.11	0.04	1.35	0.05	0.35	0.17
projected SIV	0.93	0.05	-0.07	0.06	1.00	0.07	0.00	0.07	0.66	44.22	-0.34	44.29
Agg. SIV	1.24	0.06	0.24	0.12	1.47	0.07	0.47	0.29	1.80	0.05	0.80	0.69
SIV + projected	0.93	0.05	-0.07	0.06	1.00	0.07	0.00	0.07	0.67	43.34	-0.33	43.40
SIV + Agg.	1.24	0.06	0.24	0.12	1.47	0.07	0.47	0.29	1.79	0.05	0.79	0.68
SIZ Z	1.06	0.02	0.06	0.03	1.21	0.03	0.21	0.07	1.56	0.03	0.56	0.34
					$\sigma = 2$							
OLS (TWFE)	1.47	0.02	0.47	0.24	1.73	0.02	0.73	0.56	2.03	0.02	1.03	1.08
TSLS (TWFE)	1.22	0.06	0.22	0.11	1.46	0.08	0.46	0.29	1.78	0.06	0.78	0.67
SIV	1.04	0.03	0.04	0.03	1.17	0.05	0.17	0.07	1.47	0.06	0.47	0.28
projected SIV	0.93	0.08	-0.07	0.08	1.03	0.10	0.03	0.10	1.34	0.13	0.34	0.25
Agg. SIV	1.25	0.08	0.25	0.14	1.47	0.09	0.47	0.31	1.80	0.05	0.80	0.70
SIV + projected	0.93	0.08	-0.07	0.08	1.04	0.09	0.04	0.10	1.34	0.13	0.34	0.25
SIV + Agg.	1.24	0.07	0.24	0.13	1.47	0.09	0.47	0.31	1.80	0.05	0.80	0.69
SIZ Z	1.08	0.03	0.08	0.04	1.26	0.04	0.26	0.11	1.62	0.04	0.62	0.43
					$\sigma = 4$							
OLS (TWFE)	1.62	0.04	0.62	0.42	1.94	0.04	0.94	0.92	2.30	0.03	1.30	1.72
TSLS (TWFE)	1.21	0.08	0.21	0.12	1.45	0.09	0.45	0.29	1.77	0.07	0.77	0.67
SIV	1.06	0.05	0.06	0.06	1.23	0.07	0.23	0.12	1.58	0.08	0.58	0.42
projected SIV	0.94	0.13	-0.06	0.14	1.08	0.15	0.08	0.15	1.46	0.15	0.46	0.35
Agg. SIV	1.24	0.09	0.24	0.15	1.46	0.12	0.46	0.34	1.80	0.06	0.80	0.70
SIV + projected	0.95	0.13	-0.05	0.13	1.08	0.15	0.08	0.15	1.46	0.15	0.46	0.35
SIV + Agg.	1.24	0.09	0.24	0.15	1.46	0.12	0.46	0.33	1.80	0.06	0.80	0.70
SIZ Z	1.11	0.05	0.11	0.06	1.32	0.05	0.32	0.16	1.68	0.05	0.68	0.51
					$\sigma = 8$							
OLS (TWFE)	1.82	0.06	0.82	0.74	2.23	0.07	1.23	1.58	2.67	0.07	1.67	2.87
TSLS (TWFE)	1.20	0.11	0.20	0.15	1.44	0.12	0.44	0.31	1.77	0.09	0.77	0.68
SIV	1.08	0.09	0.08	0.10	1.29	0.11	0.29	0.19	1.66	0.11	0.66	0.55
projected SIV	0.96	0.20	-0.04	0.20	1.15	0.22	0.15	0.24	1.55	0.23	0.55	0.52
Agg. SIV	1.22	0.13	0.22	0.18	1.46	0.16	0.46	0.37	1.79	0.09	0.79	0.72
SIV + projected	0.97	0.19	-0.03	0.19	1.15	0.22	0.15	0.24	1.55	0.22	0.55	0.52
SIV + Agg.	1.22	0.13	0.22	0.17	1.46	0.16	0.46	0.37	1.79	0.09	0.79	0.72
SIZ Z	1.13	0.07	0.13	0.09	1.36	0.08	0.36	0.21	1.73	0.06	0.73	0.60

A.2. Data

Turkish Statistical Institute (Turkstat) defines employment under four categories: wage-employment (60.7%), self-employment (20.3%), unpaid family worker (13.2%) and employer (5.6%). Wage-employment, or salaried employment, refers to the type of jobs that are done as an exchange for monetary or non-monetary payment. Both fixed and hourly pay are considered wage-employment under this category. The reason why we focus on salaried employment as opposed to overall employment for the empirical section of the paper is that, as suggested by Gulek (2023), wage employment and non-wage employment (self-employment, employer, or unpaid family work) are driven by different economic forces. Whereas there has to be an employer willing to hire a worker for a particular wage for that worker to have a salaried job (i.e, we can think about a labor demand curve), self-employment is an individual labor-supply decision. Natives who lose their salaried jobs due to the labor supply shock may choose to search for a salaried job while remaining unemployed, or if self-employment is a feasible alternative, may choose to remain employed. Gulek (2023) shows that transition from salaried to non-salaried jobs is an important adjustment mechanism for Turkish men but not so for Turkish women. Whereas he finds similar effects for men and women in salaried employment, he finds opposing results for non-salaried employment. He further argues that the canonical labor demand framework is more appropriate to think about wage employment (as opposed to non-wage employment) in settings where self-employment is a feasible alternative.

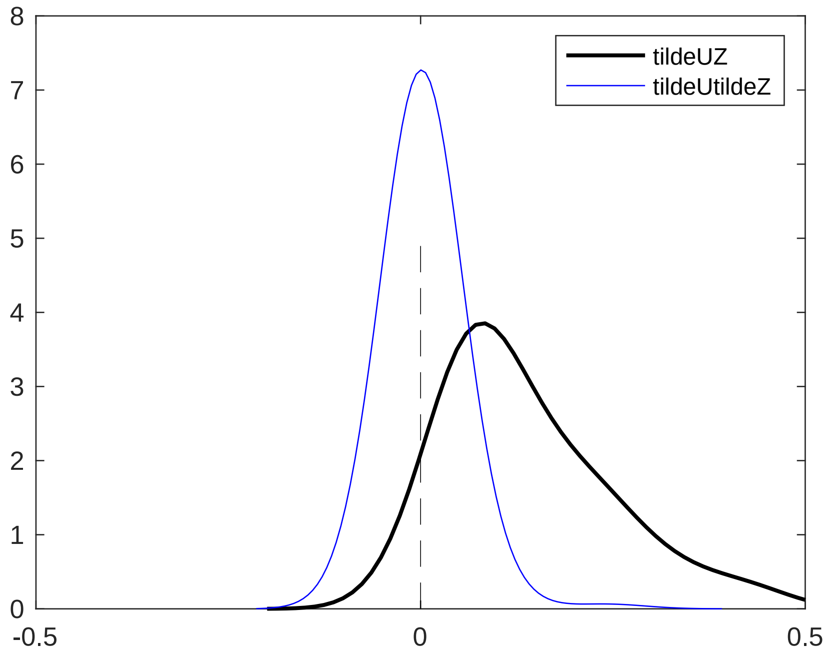
Table B.4: Educational Attainment of Syrian refugees and Natives

Educational Attainment	Syrian migrants (age 18+)	Natives (Age: 18-64)
No degree	0.21	0.12
Primary school	0.42	0.33
Secondary school	0.20	0.16
High school	0.10	0.20
Some college and above	0.08	0.19

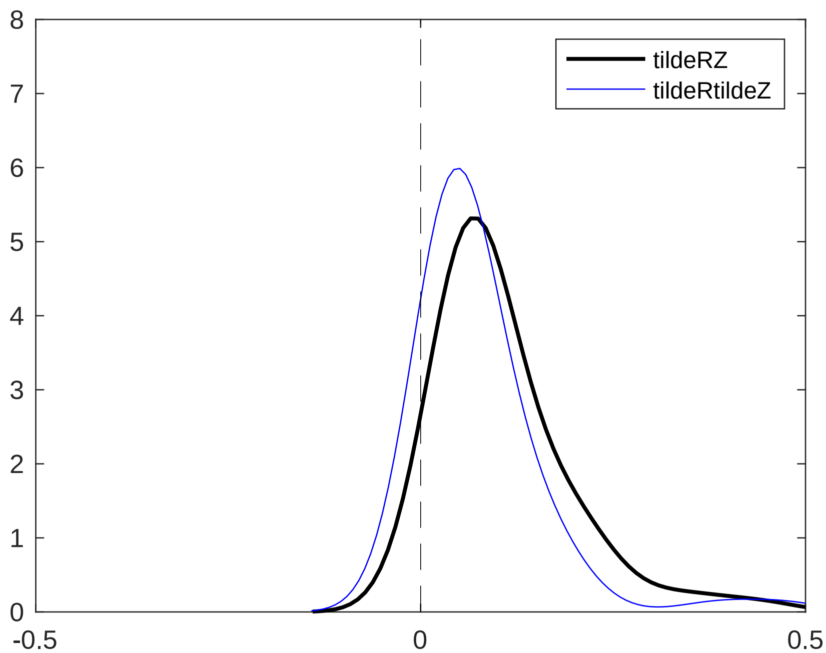
Source: Author's calculation using 2019 Household Labor Force Survey for natives, and Turkish Red Crescent and WFP (2019) for the Syrian refugees.

In the main text, we write that Syrian refugees are less educated than the Turkish natives. We show evidence for this on Table B.4. We use Turkish Household Labor force Surveys to determine the educational attainment of natives, and use livelihood surveys that are conducted on Syrian refugees to determine their educational attainment. According to these surveys, 21% of Syrian refugees in Turkey do not have any degree, 63% have at most a primary school degree, and 83% do not have a high school diploma, whereas these numbers are 12%, 45%, and 61%, respectively for natives.

Figure A.1.2: Simulation of finite sample correlations.



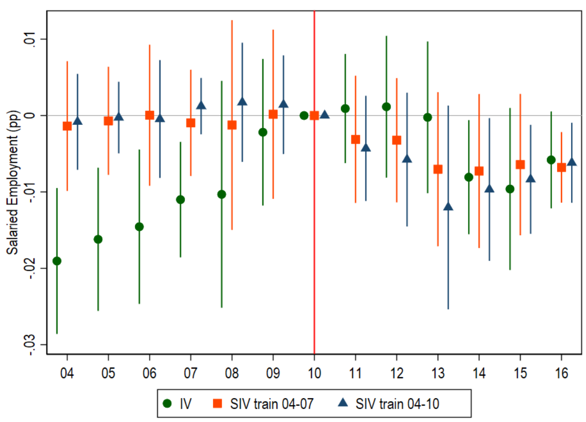
(a) Empirical average $\tilde{U}Z$ vs. $\tilde{U}\tilde{Z}$.



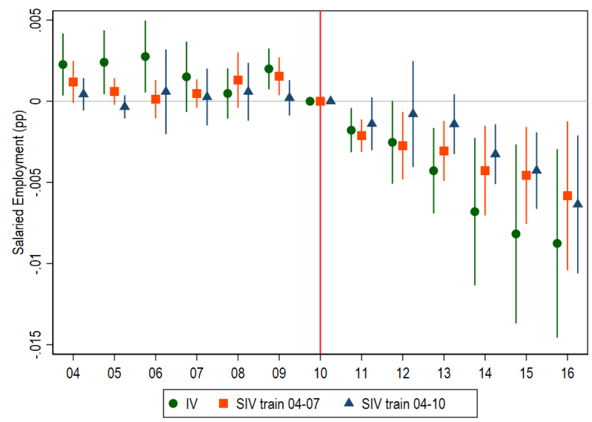
(b) Empirical average $\tilde{R}Z$ vs. $\tilde{R}\tilde{Z}$.

Notes: Panel (a) shows the distribution of $\frac{1}{JT}\sum_{it}\tilde{U}_{it}Z_{it}$ and $\frac{1}{JT}\sum_{it}\tilde{U}_{it}\tilde{Z}_{it}$ across 1000 simulations. Panel (b) shows the distribution of $\frac{1}{JT}\sum_{it}\tilde{R}_{it}Z_{it}$ and $\frac{1}{JT}\sum_{it}\tilde{R}_{it}\tilde{Z}_{it}$ across 1000 simulations. The simulation design is that of Figure A.1.1 with parameters $\rho = \rho_z = \rho_g = 0.5$ as in .

Figure A.2.3: Additional examples of IV vs SIV



(a) Men: salaried employment



(b) Women: formal salaried employment

Notes: This Figure replicates Figure 4 with an additional panel for formal women salaried employment.

A.3. Replication of Autor et al. (2013)

Table C.5: Replication of Table 3 in Autor et al. (2013)

	1990–2007 stacked first differences					
	(1)	(2)	(3)	(4)	(5)	(6)
IV	-0.75	-0.61	-0.54	-0.51	-0.56	-0.60
	(0.07)	(0.09)	(0.09)	(0.08)	(0.10)	(0.10)
SIV	-0.70	-0.59	-0.51	-0.50	-0.61	-0.63
	(0.07)	(0.10)	(0.10)	(0.09)	(0.11)	(0.10)
Controls						
Percentage of employment in manufacturing t-1	No	Yes	Yes	Yes	Yes	Yes
Percentage of college-educated population t-1	No	No	No	Yes	No	Yes
Percentage of foreign-born population t-1	No	No	No	Yes	No	Yes
Percentage of employment among women t-1	No	No	No	Yes	No	Yes
Percentage of employment in routine occupations t-1	No	No	No	No	Yes	Yes
Average offshorability index of occupations t-1	No	No	No	No	Yes	Yes
Census division dummies	No	No	Yes	Yes	Yes	Yes

Notes: The first row replicates columns 1–6 of Table 3 in ADH 2013. Row 2 shows SIV estimates. The SC weights are estimated using the manufacturing growth rates in 1970 and 1980. Dependent variable: 10 × annual change in manufacturing emp/working-age pop (in % pts). N = 1,444 (722 commuting zones × 2 time periods). All regressions include a constant and a dummy for the 2000–2007 period. Routine occupations are defined such that they account for 1/3 of US employment in 1980. The offshorability index variable is standardized to mean of 0 and standard deviation of 10 in 1980. Robust standard errors in parentheses are clustered on state. Models are weighted by start of period CZ share of national population.

A.4. Additional figure for rank effects

Figure A.3.4: Reduced-form estimates using the 1990 and 2000 shares

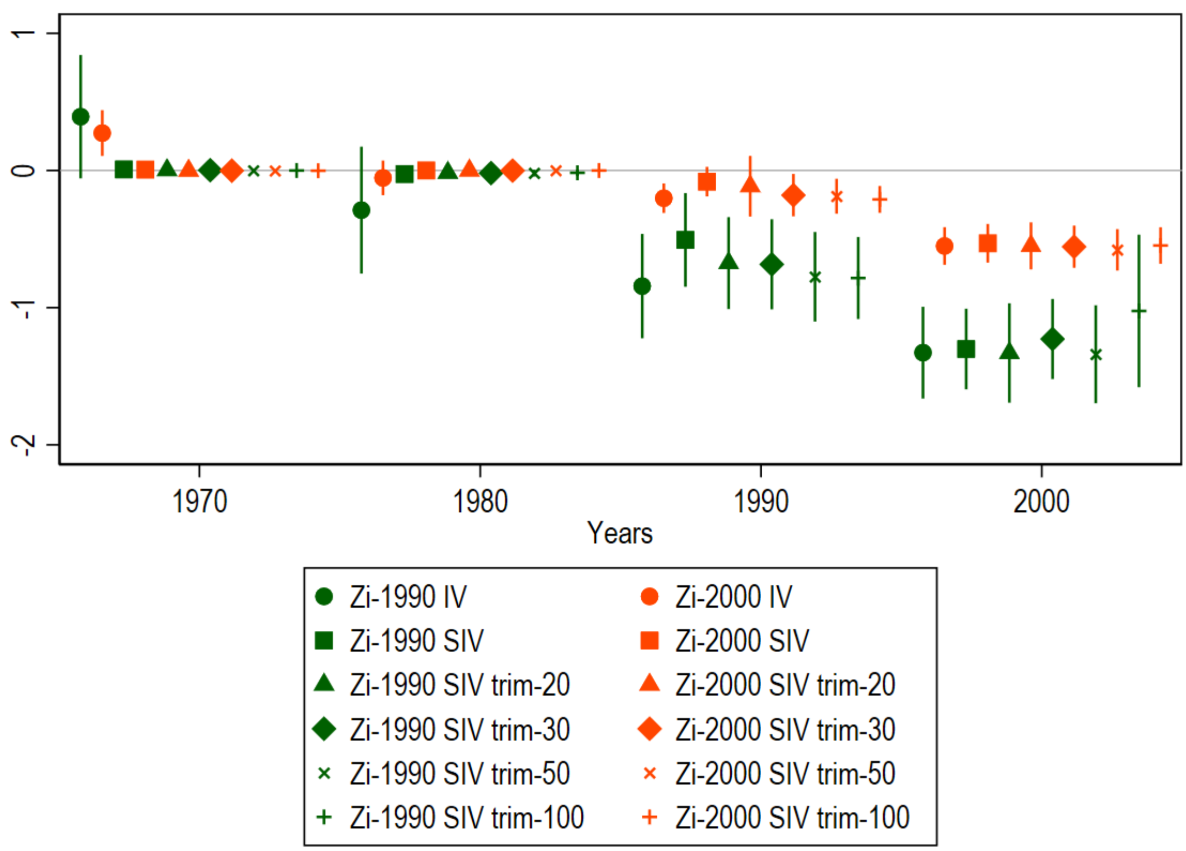


Table C.6: Replication of table 2 in Autor et al. (2013)

	1990–2000	2000-2007	1990–2007
	(1)	(2)	(3)
2SLS	-0.888	-0.718	-0.746
	(0.181)	(0.064)	(0.068)
SIV	-0.588	-0.726	-0.703
	(0.198)	(0.070)	(0.067)
SIV-trim20	-0.752	-0.763	-0.761
	(0.178)	(0.104)	(0.096)
SIV-trim30	-0.784	-0.769	-0.772
	(0.177)	(0.102)	(0.089)
SIV-trim50	-0.874	-0.807	-0.819
	(0.172)	(0.081)	(0.078)
SIV-trim100	-0.937	-0.769	-0.801
	(0.170)	(0.067)	(0.069)

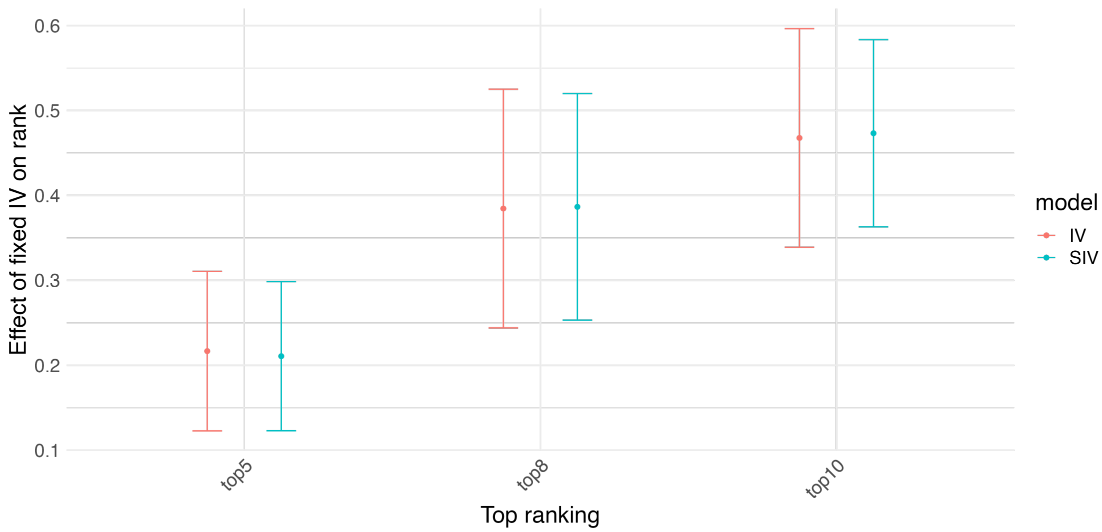
Notes: The first row replicates columns 1–3 of Table 2 in ADH 2013. In rows 2–6, we apply SIV. The SC weights are estimated using the manufacturing growth rates in 1970 and 1980. Rows 3, 4, 5, and 6 show the SIV with the donor pool trimmed to the 20, 30, 50, an 100 closest closest units to the treated unit according to the Euclidean distance, respectively.

Table C.7: Replication of table 3 in Autor et al. (2013)

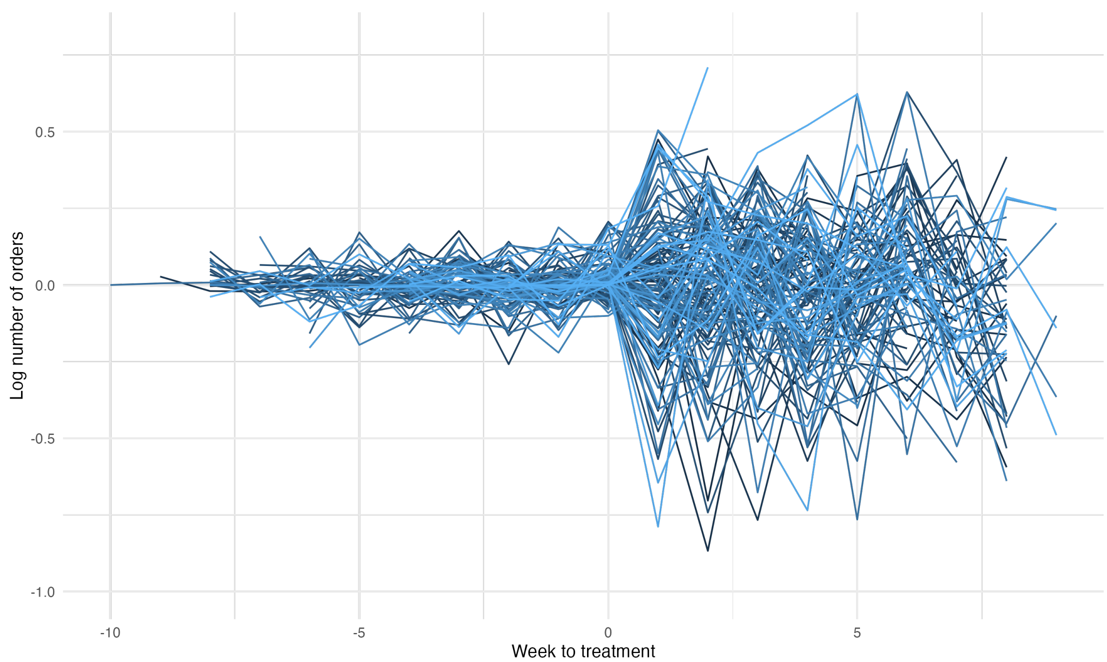
	1990–2007 stacked first differences					
	(1)	(2)	(3)	(4)	(5)	(6)
IV	-0.75	-0.61	-0.54	-0.51	-0.56	-0.60
	(0.07)	(0.09)	(0.09)	(0.08)	(0.10)	(0.10)
SIV	-0.70	-0.59	-0.51	-0.50	-0.61	-0.63
	(0.07)	(0.10)	(0.10)	(0.09)	(0.11)	(0.10)
SIV-trim20	-0.76	-0.67	-0.58	-0.57	-0.65	-0.65
	(0.10)	(0.10)	(0.07)	(0.08)	(0.08)	(0.07)
SIV-trim30	-0.77	-0.66	-0.54	-0.53	-0.59	-0.61
	(0.09)	(0.09)	(0.07)	(0.07)	(0.07)	(0.07)
SIV-trim50	-0.82	-0.74	-0.63	-0.62	-0.68	-0.70
	(0.08)	(0.09)	(0.08)	(0.08)	(0.09)	(0.09)
SIV-trim100	-0.80	-0.73	-0.63	-0.62	-0.67	-0.69
	(0.07)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)
Controls						
Percentage of employment in manufacturing t-1	No	Yes	Yes	Yes	Yes	Yes
Percentage of college-educated population t-1	No	No	No	Yes	No	Yes
Percentage of foreign-born population t-1	No	No	No	Yes	No	Yes
Percentage of employment among women t-1	No	No	No	Yes	No	Yes
Percentage of employment in routine occupations t-1	No	No	No	No	Yes	Yes
Average offshorability index of occupations t-1	No	No	No	No	Yes	Yes
Census division dummies	No	No	Yes	Yes	Yes	Yes

Notes: The first row replicates columns 1–6 of Table 3 in ADH 2013. Row 2 shows SIV estimates. The SC weights are estimated using the manufacturing growth rates in 1970 and 1980. Dependent variable: 10 × annual change in manufacturing emp/working-age pop (in % pts). N = 1,444 (722 commuting zones × 2 time periods). All regressions include a constant and a dummy for the 2000–2007 period. Routine occupations are defined such that they account for 1/3 of US employment in 1980. The offshorability index variable is standardized to mean of 0 and standard deviation of 10 in 1980. Robust standard errors in parentheses are clustered on state. Models are weighted by start of period CZ share of national population. Rows 3, 4, 5, and 6 show the SIV with the donor pool trimmed to the 20, 30, 50, an 100 closest closest units to the treated unit according to the Euclidean distance, respectively.

Figure A.4.5: First stage and SIV fit.



(a) First stage



(b) $\tilde{Y}_{it}$ fit.

Notes: Panel (a) shows the first stage regression estimates of log number of orders on R^k_{it} with week and store fixed effects. Panel (b) plots $\tilde{Y}_{it}$ (debiased log number of orders) for each producer, where we keep the 70 producers with best fit.